

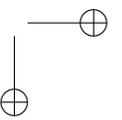
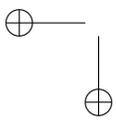
UNIVERSIDAD DISTRITAL  
FRANCISCO JOSÉ DE CALDAS

# REVISTA Ingeniería

Volumen 22 • Número 1 • ISSN 0121-750X • E-ISSN 23448393

REVISTA CIENTÍFICA CUATRIMESTRAL

# 2017



# REVISTA Ingeniería

Volumen 22 · Número 1 · Año 2017 · ISSN 0121-750X · E-ISSN 2344-8393

## REVISTA CIENTÍFICA CUATRIMESTRAL



UNIVERSIDAD DISTRITAL  
FRANCISCO JOSÉ DE CALDAS

Carrera 7 No. 40-53  
Edificio Administrativo  
Piso 4 - Facultad de Ingeniería  
Bogotá, Colombia  
Teléfono: + 57 (1) 323 93 00 ext. 2413  
Correo revista:  
revista.ing@udistrital.edu.co

<http://revistas.udistrital.edu.co/ojs/index.php/reving>

## POSTURA EDITORIAL Y AUDIENCIA

La Revista INGENIERÍA es una publicación de carácter científico con una periodicidad cuatrimestral editada por la Universidad Distrital Francisco José de Caldas. La Revista está dirigida a la comunidad académica, investigadores, egresados, sectores productivos y en general al público interesado en los temas del campo de la Ingeniería. Su principal objetivo es difundir y debatir avances en investigación y desarrollo en las diferentes áreas de la ingeniería a través de la publicación de artículos originales e inéditos, con pertinencia local o internacional.

## EDITOR

**Sergio A. Rojas, PhD.**  
Universidad Distrital Francisco José de Caldas, Colombia

## COMITÉ EDITORIAL

**Sarah Greenfield, PhD.**  
Faculty of Technology of  
Montfort University,  
Reino Unido

**Diego Cantor, PhD.**  
University of Western Ontario  
Robarts Research Institute, Canadá

**Jose Marcio Luna, PhD.**  
Perelman School of Medicine,  
University of Pennsylvania, Estados Unidos

**José Luis Villa, PhD.**  
Universidad Tecnológica de  
Bolívar, Colombia

**Luz Esperanza Bohorquez, PhD.**  
Universidad Distrital Francisco  
José de Caldas, Colombia

**Carlos Andrés Peña, PhD.**  
Institute for Information and Commu-  
nication - Haute Ecole d'Ingénierie et  
de Gestion du Canton de Vaud, Suiza

**Iván Santelices Malfanti, PhD.**  
Universidad del Bio-Bío,  
Chile

**Carlos Eduardo Moreno, PhD.**  
Universidad Nacional de  
Colombia

**Victor Hugo Grisales, PhD.**  
Universidad Nacional de  
Colombia

**Juan Carlos Figueroa, PhD.**  
Universidad Distrital Francisco  
José de Caldas, Colombia

## COMITÉ CIENTÍFICO

**Rodrigo Herrera, MSc.**  
Universidad Distrital Francisco  
José de Caldas, Colombia

**Marco Aurelio Alzate, PhD.**  
Universidad Distrital Francisco  
José de Caldas, Colombia

**Germán Méndez, PhD.**  
Universidad Distrital Francisco  
José de Caldas, Colombia

**Ana María Peña, PhD.**  
Universidad Distrital Francisco  
José de Caldas, Colombia

**Germán Jairo Hernández, PhD.**  
Universidad Nacional  
de Colombia

**Diego Pardo Ayala, PhD.**  
Agile and Dexterous Robotics  
Lab - ETHZ (Zurich)

**Edwin Rivas, PhD.**  
Universidad Distrital Francisco  
José de Caldas, Colombia

**Henry Alberto Diosa, PhD.**  
Universidad Distrital Francisco  
José de Caldas, Colombia

## DIRECTIVAS

**Carlos Javier Mosquera S.**  
Rector (E)

**Nelson Libardo Forero Chacón, PhD.**  
Director Centro de Investigación  
y Desarrollo Científico

**Roberto Ferro Escobar, PhD.**  
Decano de la Facultad de Ingeniería

## PREPARACIÓN EDITORIAL

**Carolina Suárez R., MSc.**  
Gestora Editorial

**Jenny Alexandra Jimenez, MSc.**  
Corrección de Estilo

**Julian Arcila-Forero, MSc(c).**  
Diagramación L<sup>A</sup>T<sub>E</sub>X

**Editorial Kimpres**  
Impresión

## ÁRBITROS EN ESTA EDICIÓN

**Juan Ruge Cárdenas, PhD.**  
Universidad Católica de Colombia

**Gustavo Crespo Sánchez, MSc.**  
Universidad de Cienfuegos. Cuba

**Hanwen Zhang, PhD.**  
Universidad Santo Tomás. Colombia

**Eduyn López-Santana, PhD(c).**  
Universidad Distrital FJDC. Colombia

**Giovanni Barrera Torres, PhD(c).**  
Instituto Tecnológico Metropolitano.  
Colombia

**Wilson Adarme Jaimes, PhD.**  
Universidad Nacional de Colombia

**Mauricio Hernández Sarabia, PhD(c).**  
Universidad de Ibagué. Colombia

**Eduard Gilberto, PhD (c).**  
Universidad Gilbrisco de Paula  
Santander. Colombia

**Carlos Iván Páez, PhD(c).**  
Pontificia Universidad Javeriana.  
Colombia

**Germán Chaparro Molano, PhD.**  
Universidad Sergio Arboleda.  
Colombia

**Laura Victoria Forero Vega, MSc.**  
Universidade de Sao Paulo. Brasil

**Luis Gabriel Márquez, PhD(c).**  
Universidad Pedagógica y  
Tecnológica de Colombia

**Andrea Cecilia Serge, PhD(c).**  
Universidad de Valencia. España

**Juan Pablo Gómez, PhD(c).**  
Universidad de Antioquia. Colombia

**Andrés Escobar Díaz, MSc.**  
Universidad Distrital FJDC. Colombia

**Carlos Urazán Bonells, PhD.**  
Universidad de la Salle. Colombia

**Diego Fernando García Gómez, PhD.**  
Universidad del Valle. Colombia

**Juan Carlos Figueroa, PhD.**  
Universidad Distrital FJDC. Colombia

**Gonzalo Mejía Delgado, PhD.**  
Universidad de los Andes. Colombia

**William Arnulfo Aperador, PhD.**  
Universidad Militar Nueva  
Granada. Colombia

**Iván Dario Cárdenas, PhD(c).**  
Universiteit Antwerpen. Bélgica

**Luz Edith Barba Ho, MSc.**  
Universidad del Valle. Colombia

**Karina Cela, PhD.**  
Universidad de las Fuerzas  
Armadas-ESPE. Ecuador

**Juan Miguel Becerra, PhD(c).**  
Universidad Nacional de Colombia

**Mauricio Suárez Durán, PhD(c).**  
Universidad Industrial de  
Santander. Colombia

**Omar Salazar Morales, MSc.**  
Universidad Distrital FJDC. Colombia

**Carlos Fabian Flórez, PhD(c).**  
Pontificia Universidad Javeriana.  
Colombia

**Adriana Patricia Villegas, PhD(c).**  
Universidad de Antioquia. Colombia

**Norha Ligia Posada, MSc.**  
Universidad Pontificia Bolivariana.

**Henry Diosa, PhD.**  
Universidad Distrital FJDC. Colombia

## PERMISO DE REPRODUCCIÓN

### Open access



Los textos de los artículos incluidos en esta edición pueden ser utilizados y reproducidos con fines sin ánimo de lucro y dando crédito a los autores. Está trabajo se distribuye bajo la Licencia de Creative Commons

Reconocimiento – No comercial – Sin Obra Derivada 3.0 Unported.

Ni la Revista INGENIERÍA, Editor, Comité Editorial, Comité Científico, Facultad de Ingeniería de la Universidad Distrital F.J.C., otorgan ninguna garantía, expresa o implícita, a asumen responsabilidad alguna por la exactitud, completitud o utilidad de cualquier información, aparato, producto o proceso divulgado, o que represente que su uso no infrinja derechos privados. La mención o referencia a algún producto, proceso o servicio comercial en específico, por su marca comercial, marca registrada, fabricante o cualquier otra denominación, no implica ni constituye su endoso, recomendación endoso, recomendación o favorecimiento por parte de la Revista INGENIERÍA. Los juicios y opiniones expresadas por los autores en este medio son de su responsabilidad y no establecen, reflejan o comprometen los de la Revista INGENIERÍA.

## COSTO DE PROCESAMIENTO DE ARTÍCULOS

La Revista INGENIERÍA no realiza ningún cobro por las postulaciones, evaluación y publicación de los artículos sometidos. La Universidad Distrital Francisco José de Caldas asume los gastos relacionados con el proceso de edición, gestión y publicación. Los Pares Evaluadores realizan su contribución de manera voluntaria y sin retribución económica.

## INDIZADA EN



## CARATULA

La portada está inspirada en el artículo sobre monitoreo del clima espacial desde Colombia, incluido en este número. A su vez sirve de metáfora sobre los confines inexplorados del universo de la ingeniería y su relevancia actual para el futuro del planeta.

**Idea creativa y edición gráfica:**  
Sergio A. Rojas  
**Diseño gráfico:** Julián Hernández.

## CONTENIDO

### EDITORIAL

- En Defensa del Factor  $h5^*$  Como Indicador de Impacto (\*Según Colciencias)** 05  
*In Defense of Factor  $h5^*$  as an Impact Indicator (\*according to Colciencias)*  
Sergio A. Rojas

### Ingeniería Industrial

- Revisión de la Literatura sobre el Problema de Programación de “Flow Shop” Híbrido con Máquinas Paralelas no Relacionadas** 09  
*Literature Review on the Hybrid Flow Shop Scheduling Problem with Unrelated Parallel Machines*  
Eliana Peña Tibaduiza · Edwin Alberto Garavito Hernandez · Luis Pérez Figueredo · Eslenidis Moratto Chimenty

- Modelos de Localización para Cadenas Agroalimentarias Percederas: Una Revisión al Estado del Arte** 23  
*Facility Location Models in Perishable Agri-Food Chains: a Review*  
Lizeth Andrea Sanabria Coronado · Andres Mauricio Peralta Lozano · Javier Arturo Orjuela Castro

### Ingeniería de Sistemas

- Una Mirada a la Web de los Datos. Caso de estudio: Consumo de Servicios CKAN** 46  
*A View of the Web of Data. Case Study: Use of Services CKAN*  
Jhon Francined Herrera Cubides · Paulo Alonso Gaona García · Kevin Gordillo Orjuela

### Ingeniería Eléctrica

- Análisis de Confiabilidad en Subestaciones Eléctricas Tipo Maniobra Implementando el Transformador de Tensión con Núcleo de Potencia** 65  
*Reliability Analysis in Maneuver Type Electrical Substations Implementing Tension Transformer with Power Core*  
Diego Viteri Toquica · Cesar Garzón Busto · Alexander Narváez Cubillos

### Ingeniería Electrónica

- Monitoreo del Clima Espacial Desde Colombia Mediante Radio Receptores Butterworth de Orden Superior** 83  
*Space Weather Monitoring from Colombia Through Higher Order Butterworth Radio-Receivers*  
Andrés Felipe Jaramillo Alvarado · Hamilton David Galvis Rodríguez · Edwin Andrés Quintero Salazar

### Ingeniería Ambiental

- Evaluación de los Parámetros Fisicoquímicos en un Reactor Discontinuo de Lodos Activados para el Tratamiento de Aguas con Metanol** 98  
*Evaluation of Physicochemical Parameters in Discontinuous Reactors of Activated Sludge for the Treatment of Waters with Methanol*  
Marvin Caravali · Martha L. Lote · Lena Carolina Echeverry

### Ingeniería Computacional

- Detection of Outliers and Imputing of Missing Values for Water Quality UV-VIS Absorbance Time Series** 111  
*Detección de Valores Extremos e Imputación de Valores Faltantes para la Calidad de Agua en Series de Tiempo de Absorbancia UV-VIS*  
Leonardo Plazas Nossa · Miguel A. Ávila A. · Andres Torres

- Del Operador Apertura en la Matemática Morfológica Difusa** 125  
*On the Fuzzy Opening Morphological Operator*  
Carlos Orlando Ochoa Castillo · Wilson Javier Forero Baquero

### Ingeniería Mecánica

- Estudio de la Influencia del Uso de Mezclas de Combustible Alternativo con Jet a-1 en Motores a Reacción** 140  
*Theoretical study of the influence in a jet engine using mixtures of Jet A1 and biodiesel of oil palm.*  
Luisa Fernanda Mónico Muñoz · Juan José Sandoval Sotelo · Andrés Felipe Rodríguez Chaparro

- Instrucciones para los Autores** 152



Open access



## Nota Editorial

### En defensa del factor $h5^*$ como indicador de impacto (\*según Colciencias)

El índice  $h$  fue originalmente propuesto por Hirsh [1] para medir simultáneamente la cantidad (productividad) y calidad (impacto) de la producción científica de un autor; brevemente, corresponde a la cantidad  $h$  de un total de  $n$  artículos publicados por tal autor, que hayan recibido al menos  $h$  citaciones (con lo cual  $(n - h)$  de sus artículos tendrán menos de  $h$  citaciones). De esta forma, intenta medir la visibilidad de su actividad científica compensando los sesgos que puedan originarse por un número muy grande de artículos poco relevantes o por un número muy alto de citaciones de unos pocos artículos.

Recientemente el índice  $h$  también ha sido propuesto como indicador bibliométrico alternativo para revistas científicas. Es decir, en vez de considerar un conjunto de artículos de un mismo autor, se agrupan los artículos publicados en una revista determinada y sobre ese conjunto se procede a calcular el valor  $h$  de dicha revista. Igualmente podría utilizarse para medir el impacto o pertinencia de otras unidades académicas (grupos de investigación, laboratorios, universidades). Y, aunque en principio su cálculo contempla toda la cronología de artículos asociados a una determinada unidad académica, es posible también restringirla a una ventana de observación específica, por ejemplo, aquellos publicados en los últimos diez o cinco años.

Para el caso de las revistas científicas, el valor  $h$  se puede calcular con la información de citaciones contenidas en índices bibliográficos como *Web of Science* o *Scopus*, o mediante las estadísticas académicas de *Google*, ofrecidas a través de su servicio *Google Scholar*. Este último tiene la ventaja de tener mayor cobertura por ser de acceso libre y además por incluir literatura gris en sus cálculos. De hecho *Google Scholar* reporta un escalafón de revistas científicas ordenadas según el índice  $h$ , pero restringido a fechas de publicación de cinco años atrás, factor que han denominado  $h5$  y que definen literalmente de la siguiente forma [2]:

El índice  $h5$  es el índice  $h$  de los artículos publicados en los últimos 5 años completos. Se trata del número mayor  $h$  en cuanto a que  $h$  artículos publicados entre 2011 - 2015 deben tener al menos  $h$  citas cada uno.

Nuestro interés en este último índice se debe a que el nuevo modelo de medición de revistas científicas nacionales de Colciencias-Publindex [3], lo contempla como una alternativa para definir la categoría A1, A2, B o C de una revista determinada, asociándolas a los cuartiles del factor de impacto que registre la misma, en uno de los índices mencionados anteriormente (*Web of Science* o *Scopus*) o si se encuentra fuera de ellos, a los cuartiles del índice  $h5$  de *Google Scholar*.

Citación: S. Rojas, "En defensa del factor  $h5^*$  como indicador de impacto (\*según Colciencias)". INGENIERÍA, vol. 22, no. 1, pp. 05-08, 2017.  
©Los autores; titular de derechos de reproducción Universidad Distrital Francisco José de Caldas. En línea DOI: <http://dx.doi.org/10.14483/udistrital.jour.reveng.2017.1.ne01>

Lo curioso sin embargo, es que la definición de dicho índice en el documento conceptual del modelo de Colciencias, difiere de la citada previamente. En efecto, allí definen el siguiente factor que citamos literalmente ([3], pág. 6) y que en adelante denominaremos  $h5^*$  para enfatizar que se trata de un indicador diferente:

Específicamente el  $H5$ , restringe el cálculo a las citas recibidas por una publicación científica en los últimos 5 años, independientemente de su fecha de publicación.

La diferencia no es sutil. El factor  $h5^*$  así definido, contempla todos los artículos publicados históricamente en la revista, contabilizando únicamente las citas recibidas desde cinco años atrás; en contraste, el índice  $h5$  de *Google* contempla solo los artículos (y consecuentemente las citas) de los últimos cinco años de vida de la revista. En otras palabras y a nuestro modo de ver,  $h5^*$  es un indicador del impacto de una revista combinando la vigencia e influencia de sus artículos publicados, mientras que  $h5$  únicamente refleja su vigencia.

Para ilustrar con mayor claridad el contraste, consideremos como ejemplo no una revista sino un autor emblemático: Albert Einstein<sup>1</sup>. Obviamente su factor  $h5$  es cero, puesto que no existen publicaciones de su autoría en los últimos cinco años (por supuesto hablamos de publicaciones originales, no reediciones). Ahora bien, al consultar su perfil de citas en *Google Scholar* (mostrado en la Figura 1), se observa en el recuadro que su factor  $h5^*$  es 66 (índice  $h$  desde 2011, o sea 66 de sus artículos han recibido al menos 66 citas cada uno desde el 2011). A manera de guía, el mismo Hirsch sugirió que un  $h$  de 12 caracterizaría a un profesor-investigador titular universitario promedio, un  $h$  de 20 correspondería a un científico exitoso, un  $h$  de 40 a un científico brillante digno de un Premio Nobel y uno con índice  $h$  de 60 sería un individuo extraordinariamente excepcional [1]. En este ejemplo, claramente el índice  $h5^*$  refleja acertadamente lo influyente y vigente que sigue siendo Einstein hoy en día, algo que su índice  $h5$  subestima, o más bien, desconoce.

Hasta aquí desde el punto de vista bibliométrico, el nuevo factor  $h5^*$  parece un indicador más

<sup>1</sup>Agradecemos a Marco Aurelio Alzate, nuestro Editor Asociado, por sugerirnos este ejemplo.

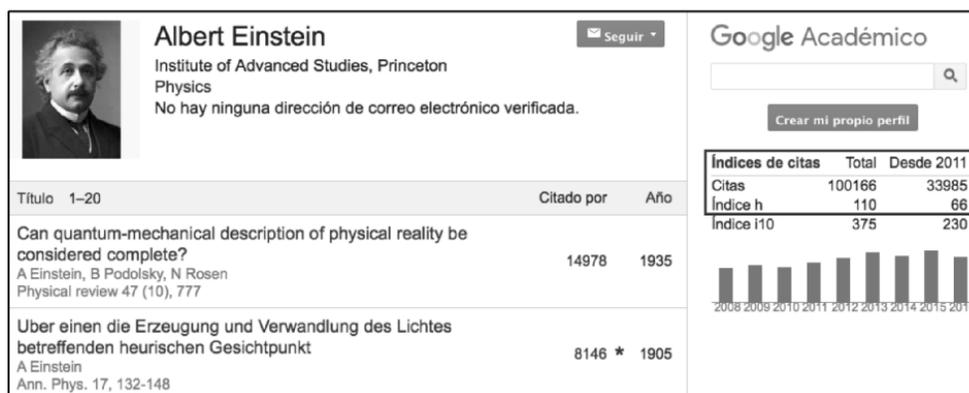


Figura 1: Fragmento del perfil de citas de A. Einstein. Tomado de *Google Scholar* (Dic.30, 2016).

apropiado para valorar el impacto científico de las revistas, pues no castiga artículos seminales que mantienen influencia más allá de la ventana de publicación de cinco años; consecuentemente nos parece adecuado asociarlo a las categorías del nuevo modelo de Publindex. A pesar de ello, lo que nos preocupa es que en la práctica Colciencias utilice una herramienta equivocada para su medición. Al parecer, según se menciona en su documento conceptual ([1], pág. 15, 16, 17), así como en las socializaciones realizadas a los editores, se pretende utilizar el software *Publish or Perish* de Harzing.com [4] que es una herramienta para consultar las estadísticas de  $h5$  (cuidado, no las de  $h5^*$ ) desde los repositorios de *Google Scholar*. Se estaría pues utilizando un instrumento que evalúa las revistas con el criterio equivocado y por lo tanto la medición resultaría distorsionada, no solamente en lo que respecta a los resultados individuales de cada revista, sino también en el cálculo de los cuartiles, que implica realizar el escalafón de la totalidad de revistas, globalmente así como por áreas de conocimiento o especialidades (ver por ejemplo la reciente caracterización de revistas A1 y A2 en el área de ingeniería realizada usando dicho software en [5]).

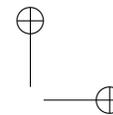
Teniendo en cuenta lo expuesto anteriormente, esta nota editorial pretende enviar dos mensajes urgentes a Colciencias y a la comunidad científica nacional:

1. Proponemos que se conserve y se respete el factor  $h5^*$  como fue definido originalmente en el documento conceptual del nuevo Publindex, para que sirva como índice alternativo a *Web of Science* o *Scopus*; a diferencia de  $h5$ , el factor  $h5^*$  combina simultáneamente productividad, influencia y vigencia en un solo indicador de impacto de una revista.
2. Sugerimos tomar las precauciones necesarias para utilizar un instrumento adecuado para la medición del índice  $h5^*$ . Desconocemos si existe en la actualidad una herramienta de software similar a *Publish or Perish* que permita calcular automáticamente los valores  $h5^*$  desde los repositorios de *Google*; por ello suponemos que para el inminente ejercicio que se avecina (Convocatoria 768 de 2016 de Colciencias), deberá realizarse mediante una inspección, depuración y recolección manual de estos valores visitando las páginas de perfiles de citación de las revistas en *Google Scholar*. Por cierto, si aún no existe tal herramienta, he allí una excelente oportunidad de desarrollo tecnológico que sería de bastante utilidad para estudios bibliométricos de la producción científica colombiana, ya sea dentro o fuera del nuevo modelo de Publindex.

**Colofón:** agradecemos y congratulamos a todos los autores, evaluadores, editores, correctores y diagramadores que participaron en la producción de este nuevo número con artículos sobre temas como el problema de programación de *flow shop*; las cadenas agroalimentarias perecederas; la Web de los datos; las subestaciones eléctricas tipo maniobra; el monitoreo del clima espacial (al cual alude nuestra carátula en esta edición); el tratamiento de aguas residuales contaminadas con metanol, los operadores morfológicos difusos y uso de mezcla de biodisel en motores a reacción. Confiamos en que será de agrado para nuestros lectores.

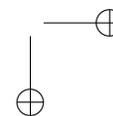
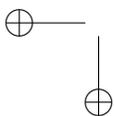
**Sergio A. Rojas, PhD.**

**Editor General Revista INGENIERÍA**



## Referencias

- [1] Hirsch, J. E. "An Index to Quantify an Individual's Scientific Research Output". *Proceedings of the National academy of Sciences of the United States of America*, 2005.
- [2] Google, Inc. *Google Scholar Metrics*. [En línea]. Disponible en: [scholar.google.com/intl/en/scholar/metrics.html](http://scholar.google.com/intl/en/scholar/metrics.html) [visitado: Dic.30/2016].
- [3] Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias). *Política Nacional Para Mejorar El Impacto De Las Publicaciones Científicas Nacionales*. 2016. [En línea]. Disponible en: [www.colciencias.gov.co/sites/default/files/upload/noticias/120816-vfpolitica\\_publindex\\_2.0\\_og\\_ao\\_miv.pdf](http://www.colciencias.gov.co/sites/default/files/upload/noticias/120816-vfpolitica_publindex_2.0_og_ao_miv.pdf) [visitado: Dic.30/2016].
- [4] Harzing, A.W. *Publish or Perish*. 2007. [En línea]. Disponible en: [www.harzing.com/pop.htm](http://www.harzing.com/pop.htm) [visitado: Dic.30/2016].
- [5] Pavas, A. "Estado del Indicador h5 en Revistas Colombianas de Ingeniería". *Ingeniería e Investigación*. 2016.



## Revisión

**Revisión de la Literatura Sobre el Problema de Programación de “Flow Shop” Híbrido con Máquinas Paralelas no Relacionadas***Literature Review on the Hybrid Flow Shop Scheduling Problem with Unrelated Parallel Machines***Eliana Peña Tibaduiza<sup>\*1</sup>, Edwin Alberto Garavito Hernandez<sup>1</sup>, Luis Pérez Figueredo<sup>1</sup>, Eslendis Moratto Chimenty<sup>1</sup>**<sup>1</sup>Universidad Industrial de Santander. Colombia

\*Correo electrónico: elmapena@uis.edu.co

Recibido: 05-04-2016. Modificado: 16-08-2016. Aceptado: 10-10-2016

## Resumen

**Contexto:** Se ha observado que el problema de *flow shop* híbrido con máquinas paralelas no relacionadas presenta escasos reportes en el ámbito académico, en comparación con el *flow shop* híbrido que considera procesadores idénticos, razón por la cual son poco conocidos los diferentes enfoques y aplicaciones que este problema puede tener.

**Método:** Se realizó una revisión de la literatura del estado del arte del problema de programación de trabajos en un *flow shop* híbrido mediante la recopilación y análisis de artículos académicos en diversas bases de datos científicas. Con este propósito se realizó una búsqueda a partir de palabras claves que definen el problema, revisando si evidencian en la descripción del problema la existencia de máquinas paralelas no relacionadas; como resultado se obtienen 50 artículos seleccionados para el presente estudio.

**Resultados:** Se realizó una clasificación del problema de acuerdo a las características del sistema productivo y se presentan los métodos de solución, restricciones y funciones objetivo comúnmente utilizadas.

**Conclusiones:** Se observa una tendencia creciente en los estudios del *flow shop* con múltiples etapas, pero pocos son los estudios basados en casos de la industria.

**Palabras clave:** “Flow shop” híbrido, máquinas paralelas no relacionadas, programación, revisión de la literatura.

**Idioma:** Español

Citación: E. Peña, E.A. Garavito, L. Pérez, E. Moratto, “Revisión de la literatura sobre el problema de programación de “flow shop” híbrido con máquinas paralelas no relacionadas” INGENIERÍA, vol. 22, no. 1, pp. 09-22, 2017.

© Los autores; titular de derechos de reproducción Universidad Distrital Francisco José de Caldas.

En línea DOI: <http://dx.doi.org/10.14483/udistrital.jour.reveng.2017.1.a01>

### Abstract

**Context:** The flow shop hybrid problem with unrelated parallel machines has been less studied in the academia compared to the flow shop hybrid with identical processors. For this reason, there are few reports about the kind of application of this problem in industries.

**Method:** A literature review of the state of the art on flow-shop scheduling problem was conducted by collecting and analyzing academic papers on several scientific databases. For this aim, a search query was constructed using keywords defining the problem and checking the inclusion of unrelated parallel machines in such definition; as a result, 50 papers were finally selected for this study.

**Results:** A classification of the problem according to the characteristics of the production system was performed, also solution methods, constraints and objective functions commonly used are presented.

**Conclusions:** An increasing trend is observed in studies of flow shop with multiple stages, but few are based on industry case-studies.

**Keywords:** Hybrid flow shop, unrelated parallel machine, literature review, scheduling.

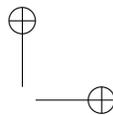
## 1. Introducción

En un sistema productivo tipo *flow shop* un conjunto de tareas debe seguir un flujo unidireccional para ser procesado a través de múltiples estaciones, donde cada estación consta de una sola máquina. En algunas fábricas, la necesidad de aumentar o equilibrar la capacidad ha llevado a la adquisición de nuevos procesadores en algunas etapas [1], los cuales pueden coexistir con los ya disponibles.

Normalmente las máquinas adquiridas tienen características diferentes de las existentes, pueden llegar a ser más rápidas, tienen un menor tiempo de alistamiento, pueden procesar más trabajos, entre otras, y son conocidas en la literatura como máquinas no relacionadas. Este tipo de sistemas es frecuente en industrias como las de fabricación de baldosas cerámicas, automotriz, metalmecánicas, de fabricación de televisores con circuito impreso a bordo y suele ser denominado *flow shop* híbrido Hybrid Flow Shop (HFS, por sus siglas en inglés) o *flow shop* flexible, Flexible Flow Shop (FFS, por sus siglas en inglés).

Tal problema ha sido del interés de numerosos investigadores desde que aparecieran los primeros trabajos en 1970, dada su complejidad y aplicación práctica. Se han planteado diferentes métodos tanto heurísticos como exactos para hallar una solución factible considerando diferentes restricciones. Tradicionalmente, los trabajos que se han realizado sobre *flow shop* híbrido han estudiado la restricción de máquinas paralelas idénticas, caso que no es muy común en las industrias manufactureras.

Es por esto que este trabajo presenta una revisión de las investigaciones desarrolladas considerando máquinas paralelas no relacionadas, con el propósito de complementar las revisiones realizadas por Ribas *et al.* [1], Morais *et al.* [2], Wenpeng Wang [3], Ruiz y Vázquez [4], en las cuales un porcentaje muy bajo de los estudios consultados corresponden a esta área. En el presente trabajo



se hace una revisión del problema determinístico, siendo este en el que se encuentran la mayor cantidad de reportes.

El contenido de este artículo se organiza como se describe a continuación: en la sección dos se presenta la definición del problema. Posteriormente, el apartado tres describe el protocolo de investigación seguido por la presentación de los resultados obtenidos (sección cuatro). Por último, se presentan las conclusiones en el apartado cinco.

## 2. Descripción del problema HFS

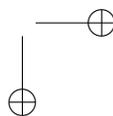
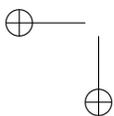
El problema de *flow shop* híbrido, Hybrid Flow Shop (HFS, por sus siglas en inglés), tiene como objetivo determinar de forma adecuada la secuencia y asignación de un conjunto de tareas a un conjunto de  $m$  máquinas para lograr optimizar una función en particular; dicha función considerada por lo general es el *makespan* o la cantidad de trabajos tardíos. En este problema se considera que los trabajos tienen el mismo orden de procesamiento y el flujo del proceso es unidireccional.

El HFS como se ha venido trabajando, no contempla aspectos importantes de la producción en las industrias reales, debido a que los estudios que generalmente se hacen no consideran parámetros y supuestos presentes en la realidad como la variación en los tiempos de alistamiento y de operación a partir de un enfoque estocástico, al igual que se excluye en la mayoría de los casos la posibilidad de un tratamiento dinámico del problema considerando las variaciones de los parámetros en periodos diferentes. De igual manera sucede con la existencia de máquinas paralelas no relacionadas. Es por esto que las revisiones que se han realizado sobre este tema son pocas y no agrupan todas las posibles variantes que puede tener este problema, hecho que conlleva a la no claridad acerca de su definición, es decir, no hay unanimidad de los conceptos, pues mientras unos autores consideran *flow shop híbrido* a aquel que tiene máquinas paralelas idénticas y *flow shop flexible* al de máquinas paralelas no relacionadas, distintos autores dicen lo contrario. Teniendo en cuenta lo dicho anteriormente se presenta una clasificación del problema de *flow shop* híbrido con máquinas paralelas no relacionadas, de acuerdo a lo reportado en la literatura por investigadores como Lin y Chen [5], Ewa Figielska [6], entre otros, considerando además el enfoque del problema como determinístico y estático.

## 3. Protocolo de investigación

El proceso de investigación inició con una revisión acerca del problema de *flow shop* híbrido con máquinas paralelas no relacionadas, teniendo en cuenta lo expuesto por Pautasso [7], quien propone diez reglas para escribir una revisión de la literatura. Asimismo, se plantean una serie de preguntas a resolver, similar a lo planteado por [8], que permiten estructurar el presente trabajo; dichas cuestiones están enfocadas a identificar cuáles son las restricciones, medidas de desempeño y métodos de solución más utilizados.

Se definió una ventana de tiempo entre los años 1995 y 2015, a su vez se construyó una ecuación de búsqueda identificando como palabras claves: “hybrid flow shop”, “flexible flow shop”,



“unrelated parallel machine”, “scheduling”. La ecuación de búsqueda se introdujo en las siguientes bases de datos: Springer Link, Elsevier, Proquest, Scopus, Web of Science e IEEE/IEE Electronic Library. Los resultados de esta consulta se resumen en la Tabla I, aclarando que en los resultados se encuentran artículos repetidos, para finalmente obtener 543 documentos diferentes.

En la selección de los documentos se consideró el hecho que no existe consenso en el medio académico sobre el nombre que recibe el problema, es decir, algunos autores consideran como *flow shop* híbrido aquel sistema que tiene máquinas paralelas idénticas y *flow shop* flexible el que presenta procesadores no relacionados, fue necesario revisar que los documentos encontrados evidenciaron en la descripción del problema la existencia de máquinas paralelas no relacionadas. Igualmente, se aplica otro filtro verificando criterios de inclusión considerando que el documento corresponda a un artículo completo; de igual manera, se estable-

ce que el contenido del documento debe presentar la descripción del modelo matemático, incluyendo la definición de la función objetivo, las restricciones y parámetros, además del método o métodos de solución planteados por los autores.

**Tabla I.** Resultados de la consulta inicial.

Base de Datos	Cantidad
Science Direct – Elsevier	252
Web of Science	194
SCOPUS	361
Proquest	280
Springer Link	205
IEEE	19

A partir de este filtro se depura la cantidad de artículos con los que se realiza la extracción de datos para continuar con la última etapa, en la cual se analizan 50 artículos, que corresponden a los que cumplieron con los criterios de inclusión. Posteriormente se realizó una síntesis sobre cada uno de los aspectos del problema buscando identificar diferentes enfoques y tendencias (restricciones, métodos de solución, publicaciones por año, etc.).

Esta revisión se enfoca en un tópico de investigación preciso, con el ánimo de brindar aspectos de interés para los investigadores de esta área, por tal razón puede considerarse una “mini-revisión”, según lo describe Velásquez [9].

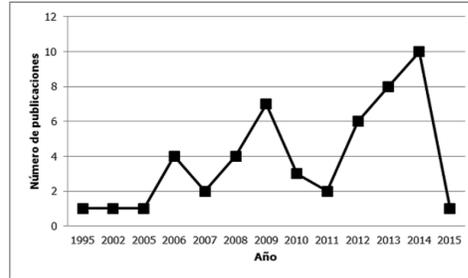
## 4. Resultados

En esta sección se muestran los resultados en cuanto a estadísticas descriptivas de la investigación realizada; es importante resaltar que al revisar los artículos seleccionados se encuentra que el 78,9% de ellos utilizan el término “híbrido” para referirse al problema en estudio, por esta razón se utilizará dicho término para la descripción de los hallazgos; es así como se observa cuáles son las tendencias de estudio del problema de *flow shop* híbrido con procesadores no relacionados.

### 4.1. Análisis de resultados

La figura 1 muestra el número de artículos publicados por año. Al analizar esta figura, se puede observar que existe una tendencia de crecimiento en las investigaciones del problema de *flow shop* híbrido con máquinas paralelas no relacionadas.

Si bien son varios los autores que han publicado sobre este problema, en los artículos encontrados para esta revisión, se observó que Ewa Figielska [6], [10]–[12] [6], [10]–[12] y Jitti Jungwattanakit [13]–[15] son los investigadores que más han incursionado en su estudio, cada uno considerando diversas restricciones: buffers de capacidad limitada y existencias de recursos no renovables por parte de Figielska y tiempos de alistamiento dependientes de la secuencia, tiempos de alistamientos dependientes de la máquina, elegibilidad de máquinas y tiempos de liberación por Jungwattanakit.



**Figura 1.** Número de publicaciones por año: detalle del número de artículos por cada año desde 1995 hasta 2015 que fueron encontrados en bases de datos científicas.

## 4.2. Clasificación del problema

Con base en lo expuesto por Linn y Zahng [15], el *HFS* puede ser clasificado en términos de la complejidad del proceso en tres categorías descritas a continuación.

### 4.2.1. HFS con dos etapas

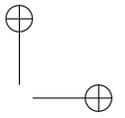
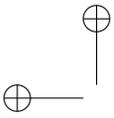
Considera la existencia de solo dos etapas de procesamiento bajo diferentes configuraciones tecnológicas. En esta categoría se pueden describir tres casos: el primero de ellos en el cual hay una sola máquina en la primera etapa y  $m$  máquinas paralelas no relacionadas en la segunda etapa; esta variación ha sido estudiada por Uetake *et al.* [17], quienes consideran el procesamiento de trabajos por lotes o familias y posteriormente por Almeder y Hartl [18], quienes lo estudian a través de un caso de aplicación en la industria metalúrgica.

Figielska [6] y Low *et al.* [19], presentan la segunda variante de este problema:  $m$  máquinas no relacionadas en la primera etapa y una máquina en la segunda. La primera considerando restricciones como preferencia de trabajo y el segundo teniendo en cuenta las restricciones de precedencia y elegibilidad de máquinas. Asimismo, la tercera y última variación:  $m$  máquinas no relacionadas en ambas etapas es investigada por Figielska [12] y Rabiee *et al.* [20].

### 4.2.2. HFS con múltiples etapas

Se han estudiado sistemas productivos que consisten de tres etapas de procesamiento en la que al menos una etapa posee múltiples máquinas paralelas no relacionadas. Sobre esta clasificación Yaurima *et al.*, realizan dos investigaciones: en 2008 [21] en industrias de fabricación de televisores con circuito impreso a bordo y luego en 2009 [20] incluyendo las zonas de almacenamiento intermedias o *buffers* con capacidad limitada.

Posteriormente Li *et al.* [23] abordan el problema considerando restricciones de grupo en la segunda etapa y proponen diez algoritmos, nueve de ellos basados en el algoritmo *Shop Partitioning (SP)* o enfoque de descomposición orientado a las etapas y uno basado en el algoritmo *Three Stage*



(TS) es decir, el enfoque holístico que considera el problema de programación.

De igual forma se ha investigado sobre trabajos que deben ser procesados en más de tres etapas. Dicho problema ha sido estudiado por Nagano y Mocellin [24] cuando lo consideran como dos *flow shop* tradicionales funcionando en paralelo. Posteriormente, el problema fue abordado por Chinyao Low [25], quien propone una heurística basada en recocido simulado para su solución.

Así como el HFS de dos etapas, el HFS de múltiples etapas presenta variaciones: la primera de ellas es el *flow shop* híbrido flexible, en el que los trabajos pueden omitir algunas de las etapas de procesamiento. Esta variante ha sido estudiada por Ruiz *et al.* [26], Zandieh *et al.* [27], Bozogiran y Logendran [28], Attar *et al.* [29], Siqueira *et al.* [30], Li *et al.* [31] y Arango *et al.* [32].

La segunda variante de esta clasificación es el *flow shop* híbrido cíclico o *flow shop* híbrido reentrante, considerada por Kim y Lee [33] cuando evalúan la tardanza total y el *makespan* y Soltani y Karimi [34], que consideran el problema sujeto a restricciones de elegibilidad de máquinas.

### 4.3. Restricciones comunes

Como se había mencionado antes, el HFS es un problema que permite modelar sistemas productivos reales si se tienen en cuenta limitaciones que se presentan a lo largo del sistema productivo, estas limitaciones generalmente tienen que ver con las tareas o grupos de tareas a ser procesadas, las máquinas y las políticas de operación. A continuación, se describen las restricciones comúnmente encontradas en la literatura:

*Tiempos de alistamiento:* Jungwattanakit *et al.* [13] en 2006 clasifican estos tiempos en dos tipos: dependientes de la máquina y dependientes de la secuencia. Un tiempo de configuración de un trabajo es dependiente de la máquina, si depende de la máquina a la cual ha sido asignado y solo ocurre con el primer trabajo asignado al procesador. Un tiempo de preparación dependiente de la secuencia es considerado entre trabajos sucesivos, es decir depende del trabajo asignado anteriormente en la máquina.

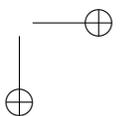
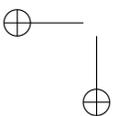
*Elegibilidad de máquinas:* no todas las máquinas existentes en cada estación son capaces de procesar todos los trabajos [35].

*Existencia de Buffers:* zonas de almacenamiento entre estaciones que pueden o no tener capacidad limitada.

*Tiempos de transporte:* consideran los tiempos de transferencia del producto entre estaciones.

*Tiempos de liberación de máquinas:* ninguna máquina puede procesar una nueva tarea hasta que no haya sido liberada. La liberación ocurre luego de transcurrir un intervalo de tiempo que incluye el procesamiento y el bloqueo, en caso que este se presente.

*No espera:* a los trabajos no les es permitido esperar entre dos máquinas sucesivas [35].



**Tabla II.** Restricciones consideradas en las publicaciones

Restricciones	Publicación donde ha sido considerada
Tiempos de alistamiento	Uetake <i>et al.</i> [17], Low [25], Ruiz y Maroto [36], Jitti <i>et al.</i> [13]–[15], Gómez [37], Ruíz <i>et al.</i> [26], Yaurima <i>et al.</i> [21], [22], Rashidi <i>et al.</i> [38], Defersha y Chen [39], Zandieh <i>et al.</i> [27], Gómez Gasquet [40], Lugo y Teixeira [41], Bozorgirad y Logendran [28], Lugo <i>et al.</i> [42], Attar <i>et al.</i> [29], Siqueira <i>et al.</i> [30], Dudas <i>et al.</i> [49], [54]
Elegibilidad de máquinas	Ruíz y Maroto [36], Ruíz <i>et al.</i> [26], Yaurima <i>et al.</i> [20], [21], Defersha y Chen [39], Lugo y Teixeira [41], Bozorgirad y Logendran [28], Lugo <i>et al.</i> [42], Siqueira <i>et al.</i> [30], Soltani y Karimi [34]
Existencia de buffers	Jenabi <i>et al.</i> [43], Figielska [6], [10], [11], Yaurima <i>et al.</i> [21], [22], Xu <i>et al.</i> [44], Almeder y Hartl [18], Lugo <i>et al.</i> [42], Soltani y Karimi [43], Li <i>et al.</i> [32], Li y Pan [45]
Tiempos de liberación de máquinas	Jitti <i>et al.</i> [14], [15], Ruíz <i>et al.</i> [26], Scholz-Reiter <i>et al.</i> [46], Zandieh <i>et al.</i> [27], Bozorgirad y Logendran [28], Lugo <i>et al.</i> [42], Siqueira <i>et al.</i> [30], Li <i>et al.</i> [23]
No espera	Rabiee <i>et al.</i> [20]
De precedencia	Ruiz <i>et al.</i> [26],
De preferencia	Parra y Mejía [46], Figielska [6], [10]–[12]

*De precedencia:* ciertos trabajos deben ser completados antes que otros puedan empezar su procesamiento [35].

*De preferencia:* se refiere a que el procesamiento de un trabajo en una máquina puede ser interrumpido e incluso posteriormente terminado en otra máquina dada una prioridad asignada a cada trabajo [35].

En la Tabla II, se presentan las restricciones que fueron consideradas en la mayoría de las publicaciones, encontrando que los tiempos de alistamiento dependientes de la secuencia y dependientes de la máquina, elegibilidad de máquinas, existencia de *buffers* y tiempos de liberación de máquinas son las restricciones más utilizadas.

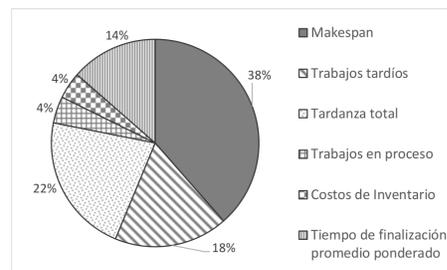
#### 4.4. Medidas de desempeño

Una de las medidas de desempeño utilizadas en mayor medida es minimizar el *makespan*, considerada por autores como: Jitti *et al.* [14], Yaurima *et al.* [22], Ruiz y Maroto [36], Gómez-Gasquet [40] y Defersha y Chen [39].

En la gran mayoría de los casos, los autores se inclinan por una única función objetivo, pero exis-



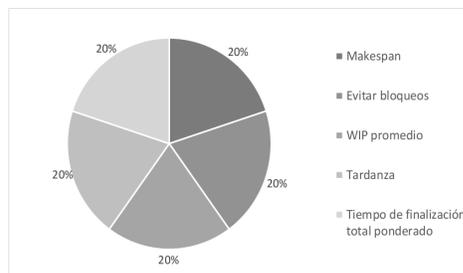
**Figura 2.** Medidas de desempeño en una sola función objetivo: muestra el porcentaje de uso de cada una de las medidas de desempeño que fueron utilizadas cuando se evaluó una función.



**Figura 3.** Medidas de desempeño en problemas bi-objetivo: detalla en que porcentaje fueron utilizadas las medidas de desempeño cuando se evaluaban dos funciones objetivo.

ten aquellos que consideran dos y más medidas de desempeño al tiempo. Tal es el caso de Dai *et al.* [48], quienes buscan minimizar el *makespan* y el consumo energético de las máquinas, Bozorgiran y Logendran [28] plantean como objetivo minimizar el tiempo de finalización total ponderado y la tardanza total ponderada.

Las figuras 2, 3 y 4, muestran cuáles son las medidas de desempeño adoptadas por los autores cuando evalúan una, dos y tres funciones objetivas respectivamente. En el caso de más de dos de funciones se encuentran trabajos como el de Dudas [48], quienes plantean minimizar el producto en proceso, Work in Process (WIP, por sus siglas en inglés) promedio, la tardanza total y el *makespan*; también Almeder y Hartl [18] presentan un caso de aplicación en la industria metalúrgica de un *flow shop* de dos etapas con existencia de un buffer limitado entre ellas; se plantearon cuatro objetivos a cumplir: utilización de las máquinas; evitar bloqueos en la máquina de la etapa uno; tiempo de finalización promedio esperado.



**Figura 4.** Medidas de desempeño en problemas multi-objetivo: muestra que cada una de las medidas de desempeño fueron evaluadas por igual en los problemas que consideraban más de dos funciones objetivas.

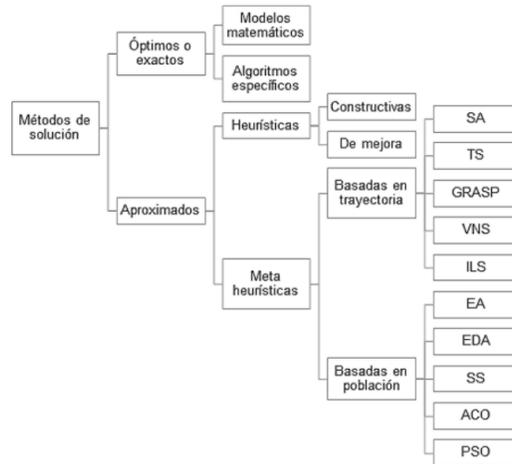
#### 4.5. Métodos de solución

Al resolver el problema de programación en un sistema productivo *flow shop* híbrido con máquinas paralelas no relacionadas muchos han sido los métodos empleados. De Fatima Morais *et al.* [2] clasifican estos métodos de solución en dos tipos (figura 5): métodos óptimos o exactos y métodos aproximados. Es de aclarar que los modelos matemáticos hacen referencia al planteamiento del problema como un modelo de programación lineal.

Ruiz y Maroto [49] definen que las heurísticas se pueden dividir en heurísticas constructivas o en heurísticas de mejora. Las heurísticas constructivas generan el paso a paso de la solución al problema, son deterministas y se basan en la mejor elección en cada iteración [51], mientras que las heurísticas de mejora toman una solución encontrada inicialmente por algún procedimiento iterativo y la modifican buscando una solución mejor de acuerdo a los objetivos definidos [2].

Finalmente, Chicano [52] en su tesis doctoral define las metaheurísticas como la combinación de diferentes métodos heurísticos a un nivel más alto para conseguir una exploración del espacio de búsqueda de forma eficiente y efectiva; a su vez clasifica las metaheurísticas en dos tipos: basadas en trayectoria y basadas en población.

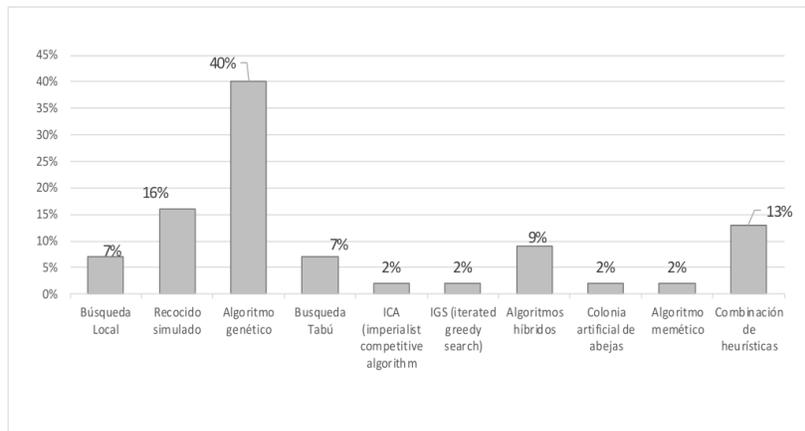
Las metaheurísticas basadas en trayectoria parten de una solución y, mediante la exploración del vecindario, van actualizando la solución inicial formando una trayectoria. Aquí pertenecen al recocido simulado. Simulated Annealing (SA, por sus siglas en inglés); búsqueda tabú, Tabu Search (TS, por sus siglas en inglés); procedimiento de búsqueda miope aleatorizado y adaptativo, Greddy



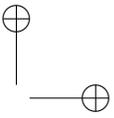
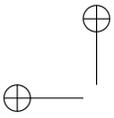
**Figura 5.** Clasificación de los métodos de solución. Presenta una adaptación de la clasificación de los métodos de solución realizada por O. Suarez [50]

Randomized Adaptive Search (GRASP, por sus siglas en inglés); búsqueda con vecindario variable, Variable Neighborhood Search (VNS, por sus siglas en inglés); búsqueda local iterada, Iterated Local Search (ILS, por sus siglas en inglés).

Por su parte, las metaheurísticas basadas en población se caracterizan por trabajar un conjunto de soluciones (llamado población) en cada iteración. En este grupo se encuentran: algoritmos evolutivos, Evolutionary Algorithm (EA, por sus siglas en inglés); algoritmos de estimación de la distribución, Estimation of Distribution Algorithm (EDA, por sus siglas en inglés); búsqueda dispersa, Scatter Search (SS, por sus siglas en inglés); optimización basada en colonia de hormi-



**Figura 6.** Métodos de solución utilizados: enseña el porcentaje de un total de 50 documentos, en que fue utilizado cada método de solución para resolver el problema dentro de las investigaciones realizadas.



gas, Ant Colony Optimization (ACO, por sus siglas en inglés); optimización basada en cúmulos de partículas, Particle Swarm Optimization (PSO, por sus siglas en inglés).

De acuerdo con el análisis realizado de las publicaciones encontradas (figura 6), se observó que el método utilizado en mayor medida es el algoritmo genético, considerado por autores como: Figielska [10], Zandieh *et al.* [27], Jun y Park [53], Lugo *et al.* [41], entre otros. Por el contrario, los menos utilizados son: el algoritmo memético, la colonia artificial de abejas, IGS e ICA, adoptados por Xu *et al.* [44], Li y Pan [45], Siqueira *et al.* [30] y Rabiee *et al.* [20], respectivamente.

## 5. Conclusiones

En este artículo se presenta una revisión de 50 publicaciones sobre el problema de *flow shop* híbrido con máquinas paralelas no relacionadas desde 1995 hasta 2015, las cuales han sido clasificadas considerando la complejidad del proceso productivo. Otros aspectos considerados son los métodos de solución planteados, las restricciones del sistema productivo y las medidas de desempeño.

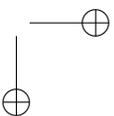
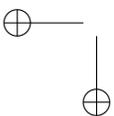
A partir de dicha revisión es posible concluir que uno de los métodos de solución comúnmente utilizados son los algoritmos genéticos y el algoritmo de recocido simulado. Por su parte, la implementación de algoritmos híbridos y metaheurísticas basadas en poblaciones como lo son los algoritmos memético o la optimización basada en colonia de hormigas (ACO) ha sido muy baja dentro de los estudios realizados, en comparación con los Algoritmos Genéticos que han sido la metaheurística más utilizada.

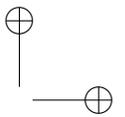
El análisis de las investigaciones realizadas en los últimos años muestra una tendencia creciente en los estudios del problema de *Flow shop* híbrido con múltiples etapas y sus variantes con un solo objetivo o múltiples objetivos, sin embargo, son pocos los casos de aplicación en las diferentes industrias los que han sido registrados en la literatura. Se han encontrado reportes de casos de fabricación de baldosas cerámicas, automotriz, metalmecánicas y de fabricación de televisores con circuito impreso a bordo. Es por esto que se evidencia un trabajo por fortalecer en llevar soluciones planteadas a las organizaciones.

Con relación a las medidas de desempeño, la minimización del *makespan* continúa siendo una de las más comunes junto con la minimización de la tardanza total y del tiempo de finalización total ponderado. El porcentaje de los trabajos que se enfocan en objetivos como la utilización eficiente de los diferentes insumos requeridos para la producción como el consumo energético o la disponibilidad de trabajadores es bajo.

En cuanto a las restricciones consideradas, son los tiempos de alistamiento dependientes de la secuencia y la existencia de *buffers* con capacidad limitada las más utilizadas. Por el contrario, restricciones como el fraccionamiento de trabajos en lotes de procesamiento, tiempos de alistamiento dependientes de la máquina y tiempos de transporte entre etapas han sido poco estudiadas.

Otro aspecto a fortalecer son las instancias disponibles en la literatura para realizar validaciones de resultados de nuevas metaheurísticas híbridas que se han venido desarrollando, pues no se en-





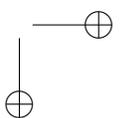
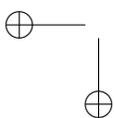
cuentran instancias con las características específicas, como por ejemplo para el caso del *flow shop* híbrido con máquinas paralelas no relacionadas con tiempos de alistamiento dependientes de la secuencia, para que así los autores puedan comparar los resultados y poder establecer la eficiencia de los algoritmos propuestos.

Dado que no se ha encontrado un gran número de artículos científicos abordando diferentes variantes al problema del *flow shop* híbrido con máquinas paralelas no relacionadas, tanto en cantidad de etapas como en sus restricciones y objetivos, se sugiere considerar restricciones como capacidad y recursos limitados, tiempos de transferencia entre etapas de procesamiento, zonas de almacenamiento entre etapas de procesamiento, entre otras, de forma que se puedan comparar los resultados y ampliar el conocimiento desarrollado hasta el momento. Asimismo, revisando trabajos previamente desarrollados y llegar a proponer modificaciones, como por ejemplo en algoritmos genéticos los operadores de cruce y mutación.

De igual forma, el problema del *flow shop* híbrido estocástico es una línea de investigación que aún no ha sido ampliamente estudiada y que puede traer consigo futuras investigaciones que lleven a solucionar problemas más cercanos a la realidad de las industrias, considerando la variabilidad de las operaciones, alistamientos, transportes y demás actividades inmersas en los procesos de fabricación. En trabajos como el de Ding *et al.* [54] se declara un interés en acercarse más a las situaciones reales de las industrias, incluyendo en su estudio la variabilidad en los tiempos de preparación y terminación.

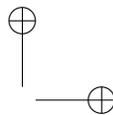
## Referencias

- [1] Imma Ribas, Rainer Leisten, y Jose. M. Framiñan, “Review and classification of hybrid flow shop scheduling problems from a production system and a solutions procedure perspective”. *Comput. Oper. Res.*, vol. 37, no. 8, 2010, pp. 1439–1454. 10
- [2] Marcia de Fatima Morais, Thays. J. Perassoli, Leandro dos Santos, Rony Peterson, y Paulo Paraíso, “Multicriteria Hybrid Flow Shop Scheduling Problem: Literature Review, Analysis, and Future Research”. *Indep. J. Manag. Prod.*, vol. 5, no. 4, pp. 2014, 1004–1031. 10, 16
- [3] Wenpeng Wang, “Review on hybrid flow shop scheduling”. *Int. Conf. Inf. Technol. Comput. Eng. Manag. Sci. ICM 2011*, vol. 4, 2011 pp. 7–10. 2011. 10
- [4] Ruben Ruiz y Jose Vázquez, “The hybrid flow shop scheduling problem”. *Eur. J. Oper. Res.*, vol. 205, no. 1, 2010, pp. 1–18. 10
- [5] James Lin y Chien Chen, “Simulation optimization approach for hybrid flow shop scheduling problem in semiconductor back-end manufacturing”. *Simul. Model. Pract. Theory*, vol. 51, 2015, pp. 100–114. 11
- [6] Ewa Figielska, “A new heuristic for scheduling the two-stage flowshop with additional resources”. *Comput. Ind. Eng.*, vol. 54, no. 4, 2008pp. 750–763. 11, 13, 15
- [7] Marco Pautasso, “Ten simple rules for writing a literature review”. *PLoS Comput. Biol.*, vol. 9, no. 7, 2013, pp. 1–4. 11
- [8] Cesar Pineda y Cesar López, “Una revisión al estado del arte de la integración de toma de decisiones en la red logística”. *Ingeniería*, vol. 18, Num. 1, pp 118-148. 11
- [9] Juan D. Velásquez, “Una Guía Corta para Escribir Revisiones Sistemáticas de Literatura Parte 1”. *Dyna*, 81. 187, 2014, pp9-10. 12
- [10] Ewa Figielska, “A genetic algorithm and a simulated annealing algorithm combined with column generation technique for solving the problem of scheduling in the hybrid flowshop with additional resources”. *Comput. Ind. Eng.*, vol. 56, no. 1, 2009, pp. 142–151. 13, 15, 18



- [11] Ewa Figielska, “Heuristic algorithms for preemptive scheduling in a two-stage hybrid flowshop with additional renewable resources at each stage”. *Comput. Ind. Eng.*, vol. 59, no. 4, 2010, pp. 509–519. 13, 15
- [12] Ewa Figielska, “A heuristic for scheduling in a two-stage hybrid flowshop with renewable resources shared among the stages”. *Eur. J. Oper. Res.*, vol. 236, no. 2, 2014, pp. 433–444. 13, 15
- [13] Jitti Jungwattanakit, Manop Reodecha, Paveena Chaovalitwongse y Frank. Werner, “Constructive and simulated annealing heuristics for hybrid flow shops with unrelated parallel machines”. in *Proceedings of the 3rd OR-CRN Operations Research Conference, Bangkok, Thailand*, 2006, pp. 110–121. 13, 14, 15
- [14] Jitti Jungwattanakit, Manop Reodecha, Paveena Chaovalitwongse, y Frank Werner, “Algorithms for flexible flow shop problems with unrelated parallel machines, setup times, and dual criteria”. *Int. J. Adv. Manuf. Technol.*, vol. 37, no. 3–4, 2008, pp. 354–370. 13, 15
- [15] Jitti Jungwattanakit, Manop Reodecha, P Chaovalitwongse, y Frank Werner, “A comparison of scheduling algorithms for flexible flow shop problems with unrelated parallel machines, setup times, and dual criteria”. *Comput. Oper. Res.*, vol. 36, no. 2, 2009, pp. 358–378. 13, 15
- [16] Richard Linn y Wei Zhang, “Hybrid flow shop scheduling: a survey”. *Comput. Ind. Eng.*, vol. 37, no. 1, 1999, pp. 57–61.
- [17] Toshifumi Uetake, Hitoshi Tsubone, y Masaaki Ohba, “A production scheduling system in a hybrid flow shop”. *Int. J. Prod. Econ.*, vol. 41, no. 1–3, 1995, pp. 395–398. 13, 15
- [18] Christian Almeder y Richard Hartl, “A metaheuristic optimization approach for a real-world stochastic flexible flow shop problem with limited buffer”. *Int. J. Prod. Econ.*, vol. 145, no. 1, 2013, pp. 88–95. 13, 15, 16
- [19] Chinyao Low, Chou Hsu, y Chwen Su, “A two-stage hybrid flowshop scheduling problem with a function constraint and unrelated alternative machines”. *Comput. Oper. Res.*, vol. 35, no. 3, 2008, pp. 845–853. 13
- [20] M. Rabiee, R. Sadeghi Rad, M. Mazinani, y R. Shafaei, “An intelligent hybrid meta-heuristic for solving a case of no-wait two-stage flexible flow shop scheduling problem with unrelated parallel machines”. *Int. J. Adv. Manuf. Technol.*, vol. 71, no. 5–8, 2014, pp. 1229–1245. 13, 15, 18
- [21] Victor Yaurima, Larisa Burtseva, y Andrei Tchernykh, “Hybrid Flowshop with Unrelated Machines , Sequence Dependent Setup Time and Availability Constraints : An Enhanced Crossover Operator for a Genetic Algorithm”. Springer verlag Berlin Heidelberg, 2008, pp. 608–617. 13, 15
- [22] Victor Yaurima, Larisa Burtseva, y Andrei Tchernykh, “Hybrid flowshop with unrelated machines, sequence-dependent setup time, availability constraints and limited buffers”. *Comput. Ind. Eng.*, vol. 56, no. 4, 2009, pp. 1452–1463. 15
- [23] Zhan Tao Li, Jianjun Liu, Qing Chen, Ning Mao, and Xiaoming Wang, “Approximation algorithms for the three-stage flexible flow shop problem with mid group constraint”. *Expert Syst. Appl.*, vol. 42, no. 7, 2015, pp. 3571–3584. 13, 15
- [24] Marcelo Seido Nagano, “Flowshops Paralelos Com Processadores Não-Relacionados”. *XXII Encontro Nacional de Engenharia de Produção*, 2002, pp. 1–8. 14
- [25] Chinyao Low, “Simulated annealing heuristic for flow shop scheduling problems with unrelated parallel machines”. *Comput. Oper. Res.*, vol. 32, no. 8, 2005, pp. 2013–2025. 14, 15
- [26] Ruben Ruiz, Funda Şerifoğlu y Thijs Urlings, “Modeling realistic hybrid flexible flowshop scheduling problems”. *Comput. Oper. Res.*, vol. 35, no. 4, 2008, pp. 1151–1175. 14, 15
- [27] M. Zandieh, E. Mozaffari y M. Gholami, “A robust genetic algorithm for scheduling realistic hybrid flexible flow line problems”. *J. Intell. Manuf.*, vol. 21, no. 6, 2010, pp. 731–743. 14, 15, 18
- [28] Mir Abbas Bozorgirad y Rasaratnam Logendran, “Hybrid flowshop scheduling problem with a bi-criteria objective and group technology assumption”. in *IIE Annual Conference. Proceedings*, 2013, p. 3004. 14, 15, 16
- [29] S. F. Attar, M. Mohammadi y R. Tavakkoli-Moghaddam, “Hybrid flexible flowshop scheduling problem with unrelated parallel machines and limited waiting times”. *Int. J. Adv. Manuf. Technol.*, vol. 68, no. 5–8, 2013, pp. 1583–1599. 14, 15
- [30] Eduardo Camargo, Sergio de Souza y Marcone Jamilson, “Um algoritmo iterated greedy search aplicado à minimização do makespan”. *Simpósio Brasileiro de pesquisa operacional*, 2013, pp. 44–55. 14, 15, 18
- [31] Zhan Tao Li, Qing Chen, Ning Mao, Xiaoming Wang y Jianjun. Liu, “Scheduling rules for two-stage flexible flow shop scheduling problem subject to tail group constraint”. *Int. J. Prod. Econ.*, vol. 146, no. 2, 2013, pp. 667–678. 14
- [32] Jaime Arango, Jaime Giraldo y Omar Castrillón, “Aplicando la Heurística TOC a la Secuenciación de Trabajos en un Flow Shop Híbrido Flexible”. *Dyna*, vol. 81, no. 2011, 2014, pp. 113–119. 14, 15
- [33] H.-W. Kim y D.-H. Lee, “Heuristic algorithms for re-entrant hybrid flow shop scheduling with unrelated parallel

- machines". *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.*, vol. 223, no. 4, 2009, pp. 433–442. 14
- [34] Abolfazl Soltani y Behrooz Karimi, "Cyclic hybrid flow shop scheduling problem with limited buffers and machine eligibility constraints". *Int. J. Adv. Manuf. Technol.*, vol. 76, no. 9–12, 2014, pp. 1739–1755. 14, 15
- [35] Omar Castrillón, Jaime Giraldo, y W. Sarache, "Técnicas inteligentes y estocásticas en scheduling". Un enfoque en la Prod. y las operaciones. Unibiblios, Bogotá, D. C., 2010. 14, 15
- [36] Ruben Ruiz y Concepción Maroto, "A genetic algorithm for hybrid flowshops with sequence dependent setup times and machine eligibility". *Eur. J. Oper. Res.*, vol. 169, no. 3, 2006, pp. 781–800. 15
- [37] Pedro Gómez, "Un nuevo algoritmo genético basado en un sistema multiagente para la programación de la producción en un taller de flujo híbrido". *Internacional Conference on Industrial Engineering & Industrial Management*, 2007, pp. 1675–1685. 15
- [38] E. Rashidi, M. Jahandar y M. Zandieh, "An improved hybrid multi-objective parallel genetic algorithm for hybrid flow shop scheduling with unrelated parallel machines". *Int. J. Adv. Manuf. Technol.*, vol. 49, no. 9–12, 2010, pp. 1129–1139. 15
- [39] Fantahun Defersha y Mingyuan Chen, "Mathematical model and parallel genetic algorithm for hybrid flexible flowshop lot streaming problem". *Int. J. Adv. Manuf. Technol.*, vol. 62, no. 1–4, 2012, pp. 249–265. 15
- [40] Pedro Gómez, Carlos Andrés y Francisco Lario, "An agent-based genetic algorithm for hybrid flowshops with sequence dependent setup times to minimise makespan". *Expert Syst. Appl.*, vol. 39, no. 9, 2012, pp. 8095–8107. 15
- [41] Pedro Luis Miranda y Rodolfo Florence Teixeira Jr, "Algoritmo genético com busca local para a programação da produção em sistemas flow-shop híbridos". *Congreso Latino-iberoamericano de Investigación Operativa*, 2012, pp. 19-29. 15, 18
- [42] Pedro Luis Miranda, Karim Martínez y Rodolfo Florence Teixeira Jr, "Um modelo de programação inteira mista para a programação da produção em flowshop híbrido com buffers limitados". *Simpósio Brasileiro de Pesquisa Operacional*, 2013, pp. 154-165. 15
- [43] M. Jenabi, S. M. T. Fatemi Ghomi, S. a. Torabi, y B. Karimi, "Two hybrid meta-heuristics for the finite horizon ELSP in flexible flow lines with 3 15
- [44] Ye Xu, Ling Wang, Gang Zhou y Shengyao Wang, "An effective shuffled frog leaping algorithm for solving hybrid flow-shop scheduling problem". *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6838 LNCS, 2011, pp. 560–567. 15, 18
- [45] Jun-qing. Li y Quan-ke Pan, "Solving the large-scale hybrid flow shop scheduling problem with limited buffers by a hybrid artificial bee colony". 15, 18
- [46] B. Scholz-Reiter, H. Rekersbrink, and M. Görge, "Dynamic flexible flow shop problems - Scheduling heuristics vs. autonomous control". *CIRP Ann. - Manuf. Technol.*, vol. 59, no. 1, 2010, pp. 465–468. 15
- [47] Javier Parra Peña y Gónzalo Mejía, "Desarrollo de un modelo para la secuenciación de trabajos en la mediana industria ladrillera de la localidad XIX", "Tesis M.S.", Unidandes, Colombia, 2006.
- [48] Min Dai, Dunbing Tang, Adriana Giret, M. Salido, y W. D. Li, "Energy-efficient scheduling for a flexible flow shop using an improved genetic- simulated annealing algorithm". *Robot. Comput. Integr. Manuf.*, vol. 29, no. 5, 2013, pp. 418–429. 16
- [49] Catarina Dudas, Marcus Frantzén y Amos HC. NG., "A synergy of multi-objective optimization and data mining for analysis of a flexible flow shop". *Robotics and Computer-Integrated Manufacturing*, vol. 27, no 4, 2011, pp. 687-695. 15, 16
- [50] Ruben Ruiz y Concepción Maroto, "Evaluación de Heurísticas para el Problema del Taller de Flujo". En *27 Congreso Nacional de Estadística e Investigación Operativa*, 2003, pp. 1–18. 17
- [51] Orlando Suarez, "Una aproximación a la heurística y metaheurísticas". *INGE@UAN, Tendencias en la Ing.*, vol. 1, no. 2, 2013, pp. 44-51. 16
- [52] Jose Francisco Chicano "Metaheurísticas e ingeniería del software". Tesis Doctoral, Lenguajes y Ciencias de la Computación, Universidad de Málaga, Málaga, España, 2007. 16
- [53] Sungbum Jun y Jinwoo Park, "A hybrid genetic algorithm for the hybrid flow shop scheduling problem with nighttime work and si multaneous work constraints: A case study from the transformer industry". *Expert Syst. Appl.*, vol. 42, no. 15, 2015, pp. 6196–6204. 18
- [54] Sahar Safaei, Reihane Naderi, Amir Sohrabi y Amin Hatami, "Scheduling of Unrelated Parallel Machines using Two Multi Objective Genetic Algorithms with Sequence-Dependent Setup Times and Precedent Constraints". *International Journal of Advanced Design and Manufacturing Technology*, vol. 8, no 4, 2016. [En línea]. Disponible en: <http://admt.iaumajlesi.ac.ir/index/index.php/me/article/view/983> 15, 19



---

**Eliana Marcela Peña Tibaduiza**

Ingeniera Industrial, Universidad Industrial de Santander; magíster en Ingeniería Industrial, Universidad Industrial de Santander; se desempeñó como docente tiempo completo en la Universidad Pontificia Bolivariana, Bucaramanga, Colombia. Actualmente se desempeña como profesor en el área de Dirección de Operaciones de la Escuela de Estudios Industriales y Empresariales en la Universidad Industrial de Santander de Bucaramanga; pertenece como investigador al Grupo de Organización y Optimización de Sistemas Productivos, Administrativos y Logísticos - OPALO y es coordinadora del Grupo de Aplicación de Lúdicas como Estrategia de Enseñanza-Aprendizaje -GALEA.  
Correo electrónico: elmapena@uis.edu.co

---

**Edwin Alberto Garavito Hernández**

Ingeniero Industrial Universidad Industrial de Santander; especialización en Docencia Universitaria y en Gerencia de la Producción y Mejoramiento continuo, Universidad Industrial de Santander; magíster (C) en Ingeniería Industrial, Universidad de Puerto Rico; actualmente se desempeña como profesor en las áreas de Simulación y Diseño de Sistemas Productivos de la Escuela de Estudios Industriales y Empresariales en la Universidad Industrial de Santander; investigador al Grupo de Organización y Optimización de Sistemas Productivos, Administrativos y Logísticos – OPALO.  
Correo electrónico: garavito@uis.edu.co

---

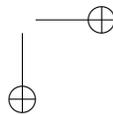
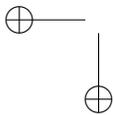
**Luis Eduardo Pérez Figueredo**

Ingeniero Industrial de la Universidad Industrial de Santander; profesional responsable de planeación del Instituto Municipal de Empleo y Fomento Empresarial De Bucaramanga.  
Correo electrónico: lpeduardop@gmail.com

---

**Eslendis Moratto Chimenty**

Ingeniera Industrial, Universidad Industrial de Santander. Correo electrónico: esle1993@gmail.com



## Revisión

**Modelos de Localización para Cadenas Agroalimentarias  
Perecederas: una Revisión al Estado del Arte***Facility Location Models in Perishable Agri-Food Chains: a Review***Lizeth Andrea Sanabria Coronado<sup>1</sup>, Andres Mauricio Peralta Lozano<sup>1</sup>,  
Javier Arturo Orjuela Castro\*<sup>1</sup>**<sup>1</sup>Universidad Distrital Francisco José de Caldas Facultad de Ingeniería.

\*Correspondencia: jorjuela@udistrital.edu.co

Recibido: 03/05/2016. Modificado: 17/08/2016. Aceptado: 10/10/2016.

## Resumen

**Contexto:** El problema de la localización de instalaciones de cadenas de suministros de productos agrícolas perecederos no ha sido suficientemente estudiado, sin embargo, el tema en los últimos años ha tomado interés debido a la preocupación por el consumo de alimentos frescos y saludables, aspecto relevante en países con potencial productor en alimentos hortofrutícolas.

**Método:** Se realizó una revisión sistemática del estado del arte del problema de localización de instalaciones para alimentos agrícolas con énfasis en perecederos; se consultaron las bases de datos: Science Direct, IEEE Xplore, SpringerLink, JSTOR, Scopus y Google Scholar, de 135 artículos de localización en alimentos, finalmente se seleccionaron 31 donde contemplan los modelos de localización en cadenas de suministro de alimentos, con énfasis en perecederos, basados en el criterio de la última década 2004-2015.

**Resultados:** En la revisión se propone una tipología para la clasificación de variables contempladas en los modelos, se establecieron las variables no contempladas en los modelos, pertinentes para países con las características de Colombia. Se identificaron las funciones objetivo, comúnmente usadas, así como las restricciones específicas para las Cadena de Suministro Alimentaria (CSA) y los drivers logísticos más utilizados en estos modelos. Se identifican líneas de futuras investigación, respecto a variables a incluir en modelos de localización para CSA de perecederos (CSAP) en contexto de países montañosos.

**Conclusiones:** Decisiones estratégicas relevantes como la localización de instalaciones son de importancia durante el diseño de la CSA y de su red de distribución, dado el carácter perecedero de los alimentos frescos. Sin embargo, el tema ha sido poco tratado debido especialmente a la dificultad para el modelamiento y al número de variables a considerar en las CSAP. Se requieren modelos que contemplen la perecibilidad, las pérdidas y el comportamiento estocástico de algunas variables. Países montañosos como Colombia abren puertas a investigaciones que modelen el cambio en los climas, pisos térmicos, condiciones de viaje y su incidencia en la perecibilidad de los alimentos.

**Palabras clave:** Cadena de suministro agro-alimentaria, localización de instalaciones, productos perecederos.

**Idioma:** Español

Citación: L. Sanabria, A. M. Peralta, J. A. Orjuela, "Modelos de Localización para Cadenas Agroalimentarias Perecederas: una Revisión al Estado del Arte" INGENIERÍA, vol. 22, no. 1, pp. 23-45, 2017.

© Los autores; titular de derechos de reproducción Universidad Distrital Francisco José de Caldas. En línea DOI: <http://dx.doi.org/10.14483/udistrital.jour.reveng.2017.1.a02>

### Abstract

**Context:** The problem of locating facilities in supply chains of perishable agricultural products has not been widely addressed, however, this issue in recent years has taken great interest because of concerns about the consumption of fresh and healthy food, important aspect in countries with fruits food producer potential.

**Method:** A systematic review of the state of the art on facility location problems for perishable agricultural products was conducted. The following databases were searched: Science Direct, IEEE Xplore, SpringerLink, JSTOR, Scopus and Google Scholar. From an initial collection of 135 papers obtained with a generic search query for food location, 31 were finally selected for review focusing on food supply chain location model, within a publication date ranging from 2004 to 2015.

**Results:** Based on the review, a typology of variables considered and not considered in models pertinent to conditions relevant to countries such as Colombia, is proposed. Objective functions commonly used, specific restrictions for Food Supply Chain (FSC) and logistical drivers most commonly considered in these models were also identified. Future research, regarding variables to be included in FSC models location for perishables in context of mountainous countries were identified.

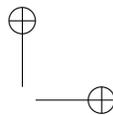
**Conclusions:** Strategic decisions such as the location of facilities are of great importance for the design of the fresh food supply chain and its network, given the perishability. However, the issue has been poorly evaluated especially due to the difficulty of modeling and the large number of variables to consider in this type of chain. Models that include the perishability, losses and stochastic behavior of some variables are required. Colombia can benefit from more research in location models of perishable food, particularly the impact on perishability due to its particular conditions of great variety of climates, mountainous landscapes along with a limited transport and road infrastructure.

**Keywords:** Location of facilities, supply chain Agro-Food, perishable products.

## 1. Introducción

La cadena de suministro agroalimentaria (CSA) es una red de organizaciones que trabajan juntas, en diferentes procesos y actividades, con el fin de llevar los productos agrícolas de la granja a la mesa y satisfacer las demandas de los consumidores [1], [2], está conformada por productores, mayoristas, distribuidores, corredores, empresas de servicios, empresas al por menor y los consumidores [3], [4]. Se diferencia de otras CS por la importancia que desempeñan factores como la calidad y la seguridad alimentaria [5], [6]; según la FAO (FAO, 2011) más del 50 % de las frutas producidas en América Latina se pierden o se desperdician antes de llegar a su destino final.

Un proceso complicado en la CSA es la distribución de alimentos perecederos con calidad [4]. Un producto se dice que es perecedero si alguna de sus características presenta deterioro con respecto a los requisitos de los productores o del cliente [7], [8]. Los productos perecederos comienzan a deteriorarse desde el momento en que se producen, parte de su pérdida se debe al paso del tiempo [9], a los cambios frecuentes de temperatura y al inadecuado mantenimiento de la calidad a lo largo de la cadena [4], [9].



Las CSA de frescos y perecederos se caracterizan por una vida corta de los alimentos, el transporte rápido [4], vida útil limitada, variabilidad en la demanda y precios; esto hace las cadenas más complejas y difíciles de manejar [ 2], ya que solamente pueden ser trasladados en la CSA durante un lapso de tiempo máximo, luego de ser cosechados [11]. En este tipo de productos una correcta gestión logística permite una reducción de los costos logísticos y ofrecer mejores alimentos a los consumidores, con precios más bajos [10] [7].

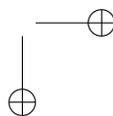
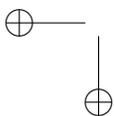
La preocupación por la calidad de los alimentos, la salud y el medio ambiente por parte de los consumidores se ha incrementado, para ello se busca que el tiempo de permanencia y envío entre las instalaciones sea el menor posible, por lo tanto, la localización de las instalaciones en la CSA contribuye con este propósito al mantener el alimento fresco [11]. El problema de localización de instalaciones (PLI) agrícolas tiene relevancia en los países en desarrollo, donde la producción agrícola no se encuentra concentrada y el tamaño medio de las explotaciones es pequeño comparado con el número de agricultores, el PLI en la agricultura es de un tamaño considerable magnitud, consideran la ubicación de múltiples instalaciones, heterogeneidad y limitaciones de capacidad [12]], en los últimos años se ha incluido el medio ambiente y la competitividad de la CS al problema [13], [14].

En los problemas de localización los centros de distribución son nodos importantes de los sistemas logísticos, [15], [16] su ubicación determina las restricciones de las redes y afecta directamente la eficiencia y la eficacia de la operación misma [17], [18], [19]. A diferencia de un producto industrial, el producto agrícola perecedero tiene un menor valor, mayor nivel de consumo, un proceso de distribución más complejo, una vida útil corta y es más difícil de gestionar en el almacén. Estos factores se ven reflejados en mayores costos logísticos que derivan en un aumento en el precio de venta [9], estos elementos derivan en la necesidad de establecer variables y parámetros específicos para el PLI en CSA de perecederos.

Con base en lo anterior este artículo se centra en la revisión de la literatura de los trabajos relacionados con la localización de instalaciones específicamente en las CSA perecederas, con el objetivo de identificar la forma como se ha abordado en el PLI el carácter perecedero de los alimentos frescos e identificar las formas de solución y posibles aplicaciones en mercados emergentes. La motivación del mismo se centra en el contexto colombiano, donde existen diversas problemáticas en la distribución logística para este tipo de cadenas, relacionadas con la gestión del proceso, el ruteo y la localización de los nodos o instalaciones dentro de la red [20].

Respecto a los nodos de transferencia, existen dos dificultades, la primera atañe a la ubicación de los centros de acopio donde ingresa los alimentos frescos que provienen de los municipios, se alista y acondiciona para ser enviada a mayoristas, minoristas y plantas procesadoras [21]. La segunda dificultad está relacionada con los puertos, debido a su ubicación en la Costa Caribe y Pacífica, los trayectos de envío terrestres son extensos, lo cual genera daños en el producto.

A pesar que en el país se han invertido recursos en la construcción de centros de acopio a nivel municipal y departamental, los resultados no han sido los esperados, dado que los únicos centros de acopio que funcionan de manera aceptable son los que constituyen las grandes Centrales de Abastos de las principales ciudades del país [22]. En ese sentido la localización de instalaciones jugará un papel fundamental en el mejoramiento del desempeño logístico de la CSA de perecederos, ya



que el diseño de la red de distribución es una decisión crítica en la gestión de operaciones por los costos, tiempo y calidad de alimento [23].

A partir de la revisión de la literatura, el artículo busca responder a las preguntas: ¿por qué es importante la ubicación de instalaciones en países con alto potencial en CSA de perecederos como Colombia?, ¿cuáles tipos de modelos han sido aplicados para la ubicación de instalaciones en las CSA de perecederos?, ¿Cual formulación de modelos se utilizan con mayor frecuencia en CSA de perecederos?, ¿cuáles son los métodos de solución aplicados al PLI de CSA de perecederos?, ¿serán aplicables estos modelos y métodos a las características de la CSA de Colombia?

El presente artículo se desarrolla de la siguiente manera: en la Sección 2 se presenta la metodología del levantamiento, selección y clasificación de la información. En la Sección 3, la definición y clasificación de la localización de instalaciones en CSA y en la Sección 4 la discusión y conclusiones.

## 2. Metodología

Para el desarrollo del artículo se consultaron las siguientes bases de datos: Science Direct, IEEE Xplore, Springer Link, JSTOR, Scopus y Google Scholar. Las palabras clave de búsqueda se presentan en la Tabla I. Las búsquedas se realizaron en Title, Abstract y Keywords en cada una de las bases de datos.

La Figura 1 presenta el consolidado de las bases de datos Scopus, IEEE Xplore, SpringerLink y JSTOR, de la cantidad de artículos publicados utilizando las ecuaciones de búsqueda: Facility Location Agri-Food y Facility Location Perishable, en el entorno de ingenierías, localización, logística y cadenas de suministro. Se observa que el número de publicaciones en productos perecedero han sido pocas a través de los años.

Luego de realizar una revisión exhaustiva de la literatura se identificaron los tipos de PLI usados con mayor frecuencia en la CSA, se construyó una visión general y se completó con la búsqueda de problemas aplicados específicamente a productos perecederos, también se identificaron los tipos de formulación empleados y los métodos de solución aplicados.

Tabla I. Palabras claves de búsqueda	
Categoría	Palabra de busqueda
Campo de aplicación	Fruit Supply Chain, Supply Chain Distribution Models, Logistics, Agri Food Supply Chain, Food Supply Chain
Palabras relacionadas con localizacion	Facility, Location, Allocation
Ecuaciones de busqueda de localizacion en general	Facility Location Logistics, Facility Location Supply Chain, Problem Facility Location, Facility Location
Ecuaciones de busqueda de localizacion de perecederos	Facility Location Perishable, Problem Facility Location Perishable, Facility Location Agri Food, Facility Location food, Facility Location Logistics

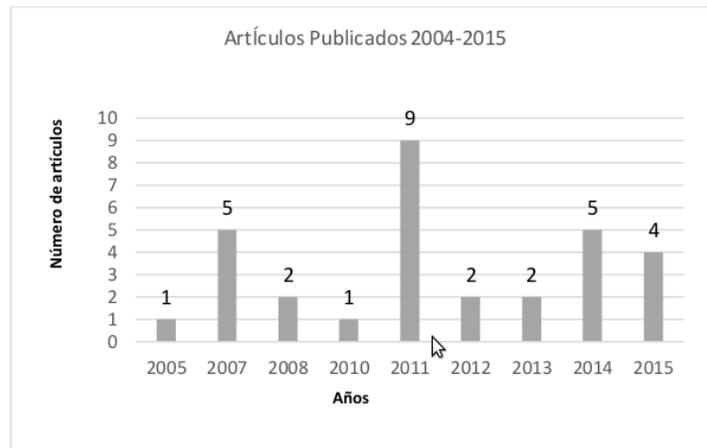


Figura 1. Artículos Publicados

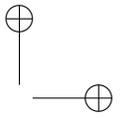
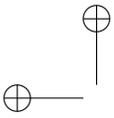
### 3. Localización de instalaciones en cadenas agroalimentarias

La localización de instalaciones es: “el proceso de elegir un lugar geográfico entre varios para realizar las operaciones de determinado actor” [24], se relaciona con los centros de distribución (CD), fábricas y centros de acopio (CA), entre otros, que deben ser seleccionados y asignados para ser atendidos en la cadena, por los diferentes actores [25]. Los problemas de localización se refieren al modelado, formulación y solución para la ubicación de instalaciones [26]. El campo de la teoría de la localización y el modelado tiene sus raíces en la primera mitad del siglo 20 con los trabajos desarrollados por Weber [27], Hotelling [28], Christaller [29] y Lösch [30]. En la actualidad sigue siendo un campo activo [31].

Las decisiones de localización son un aspecto crítico para garantizar que una CS sea eficiente, debido a que una mala ubicación dará lugar a un exceso de costos reflejados durante la vida útil de las instalaciones [32]. El establecimiento de nuevas instalaciones, la ampliación, reducción y reubicación de instalaciones hacen parte del nivel estratégico de decisión en la CS, suelen ser proyectos a largo plazo e implican actividades que consumen tiempo y capital [33], [34].

La capacidad, el nivel de comercialización y la calidad de los productos dependen en parte de la ubicación de las instalaciones y la relación existente con los demás agentes en la red logística [35]. La ubicación de instalaciones se combina frecuentemente con decisiones de inventario y producción [33]. Este tipo de decisiones tienen en cuenta los costos, rentabilidad, tiempos de respuesta, cercanía a determinados lugares de acuerdo a las características del sector o cadena evaluados [24].

Los modelos de ubicación y asignación comprenden formulaciones que varían en complejidad desde un modelo lineal simple hasta modelos probabilísticos no lineales, cambian ampliamente en cuanto a los supuestos fundamentales, complejidad matemática y rendimiento computacional. Los algoritmos de solución incluyen, entre otros, la búsqueda local y enfoques basados en programación matemática [35].



### 3.1. Tipos de problemas de localización

En esta sección se presenta una revisión al estado del arte de los tipos de problemas de localización enfocado en productos perecederos, la notación utilizada se presenta en Anexo I, la cual se basa en ReVelle *et al.* [31]. El primer problema de ubicación de instalaciones se plantea por el matemático Torricelli, es aplicado a tres clientes y una instalación, con el objetivo de minimizar la suma de las distancias euclidianas entre ellos. Los modelos de ubicación y asignación comprenden formulaciones que varían en complejidad desde un modelo lineal simple hasta los modelos probabilísticos no lineales, cambian ampliamente en cuanto a los supuestos fundamentales, complejidad matemática y rendimiento computacional. Los modelos simples intentan localizar una sola instalación y los complejos podrían localizar “p” instalaciones. Los problemas pueden ser capacitados o no capacitados. Las limitaciones de capacidad permiten que la demanda de un cliente se pueda satisfacer por varias instalaciones, lo que deriva en un problema más complejo [26].

En los modelos capacitados existe un conjunto de ubicaciones potenciales para las plantas, que suministrarán a un grupo de clientes de bienes, el problema busca encontrar el subconjunto de instalaciones o plantas que minimice los costos totales fijos y de transporte, de tal manera que la demanda de todos los clientes pueda ser satisfecha sin violar las restricciones de capacidad, cada planta tiene una capacidad potencial o un nivel máximo de servicio [36].

Según Owen [37] los modelos de localización se pueden dividir en estocásticos, determinísticos, dinámicos y continuos. A continuación serán descritos los tipos de problemas de localización utilizados para la CSA con énfasis en perecederos, se presenta la formulación clásica de cada tipo y se especifica cuáles son las variables y restricciones que se han incluido para este tipo de cadenas. Son presentados los problemas P-Mediana, P-Centro, P-Cobertura, Hubs, Multi-producto, Servicios Múltiples y capacitados.

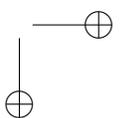
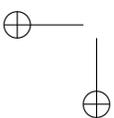
#### 3.1.1. Problemas P-Mediana

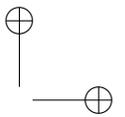
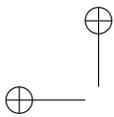
Los problemas P-Mediana son los más simples en la categoría de los modelos discretos, tienen como objetivo principal la selección de la ubicación de las instalaciones, minimizando las distancias totales ponderadas o los costes para suministrar las demandas del cliente, asumen el mismo costo de localización en cada sitio candidato [33], [35]. Se decide sobre la ubicación las instalaciones que deben atender a cada nodo que demanda, con un objetivo basado en costos o en ganancias. El problema P-Mediana toma como entradas las demandas (o pesos) de cada nodo, las distancias entre estos, la posible ubicación de la instalación, el número de instalaciones a ser ubicadas, la estructura de costos, las capacidades y las economías de escala. [31]. En la solución de algunos problemas las demandas se pueden dividir entre las diferentes instalaciones.

A continuación se presenta el modelo general de los problemas P-Mediana [31]:

$$\text{F.O} \quad \text{Minimizar} \quad \sum_{j \in J} \sum_{i \in I} w_i d_{ij} y_{ij} \quad (1)$$

Restricciones:





$$\sum_{j \in J} y_{ij} = 1 \quad \forall i \in I, \quad (2)$$

$$y_{ij} - x_j \leq 0 \quad \forall i \in I, \forall j \in J \quad (3)$$

$$\sum_{j \in J} X_j = p, \quad (4)$$

$$x_j \in \{0, 1\} \quad \forall j \in J, \quad (5)$$

$$y_{ij} \in \{0, 1\} \quad \forall i \in I, \forall j \in J \quad (6)$$

En la función objetivo (1) se minimiza la distancia media ponderada de la demanda. La restricción (2) asegura que se asigne a cada nodo de demanda una instalación candidata, mientras que la (3) controla que solo se asigne aquellas instalaciones localizadas. Por su parte la restricción (4) controla que se abra el número de instalaciones requeridas y las restricciones (5) y (6) definen las variables de decisión como binarias.

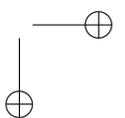
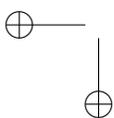
En la CSA los problemas P-Mediana se utilizan para la ubicación de un solo tipo de instalación principalmente centros de distribución [38], [39], [7], dado el tipo de alimento que se comercializa en estas cadenas es importante la localización de plazas y mercados para los agricultores [11]. Al igual que en los modelos P-Mediana generales, los modelos aplicados a productos agrícolas tienen como objetivo principal la minimización del costo total, que incluye entre otros los costos logísticos, de transporte, de construcción y de operación en cada una de las instalaciones. Los pocos trabajos que incluyen el carácter perecedero lo hacen a través de costos de penalización, costos de pérdida y costo de desperdicio, estos modelos añaden la temporalidad y el espacio debido al carácter perecedero de los productos agrícolas, la pérdida de producto y tasas de descomposición [7], [38], [39], [11]. Para su aplicación en alimentos perecederos, los autores Xiaohui [39] y Zhang [2] tuvieron en cuenta el carácter perecedero como una restricción adicional dentro del modelo. La logística de frutas y verduras es diferente a la de otro tipo de productos, debido a la facilidad que tienen para perder sus propiedades, genera un alto nivel de desperdicio y un mayor costo logístico [39]. En ese sentido Zhang [7] desarrolla una función de pérdida penalizada con un costo, la cual está relacionada con la velocidad y la distancia entre la instalación y el punto de demanda.

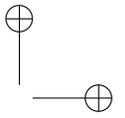
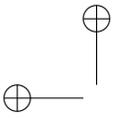
$$W_{ij}(0) - W_{ij}(t_{ij}) = d_j(e^{\theta t_{ij}} - 1) \quad (7)$$

$$f d_{ij} = c d_j (e^{\theta t_{ij}} - 1) \quad (8)$$

$$F d_{ij} = \sum_{j=1}^n c d_j (e^{\theta t_{ij}} - 1) \quad (9)$$

La ecuación (7) calcula la cantidad de producto perdido en la transferencia desde el punto  $i$  hasta el punto  $j$  y las ecuaciones (8) y (9) obtienen el valor de la pérdida total en el sistema evaluado. Por su parte Xiaohui [39] realiza un análisis similar al de Zhang con respecto al cálculo del costo





de perecibilidad, pero combina variables espacio y tiempo con el fin de eliminar la discrepancia existente entre la ubicación del consumidor y el productor. El autor presenta la ecuación (10) para calcular el costo del carácter perecedero.

$$cd_i(e^{\theta\alpha d_{ij}} - 1) \quad (10)$$

Tong [11] quien también trabajo el problema P-Mediana en productos perecederos, solamente consideró el crecimiento poblacional como criterio para la localización de las instalaciones, sin incluir la perecibilidad de los productos.

### 3.1.2. Problemas de cobertura

En este tipo de modelos se busca suplir la demanda de diferentes nodos, con una distancia o tiempo máximo aceptable, es decir un plazo determinado. Reduciendo el costo total de la ubicación de instalaciones y maximizando el nivel de servicio hacia cualquiera de los clientes en el problema. La demanda de un nodo o cliente se dice que está cubierta, si la distancia o el tiempo entre este y su instalación más cercana no es mayor a un valor especificado previamente [26]. La función objetivo y la primera restricción de los problemas P-Cobertura difieren del modelo P-Mediana [31]:

F.O:

$$\text{Maximizar } \sum_{i \in I} w_i z_i \quad (11)$$

Restricciones:

$$z_i - \sum_{i \in I} x_j \leq 0 \quad \forall_i \in I, \quad (12)$$

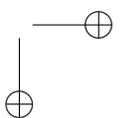
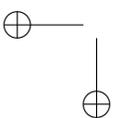
La función objetivo (11) maximiza el número de demandas cubiertas, la restricción (12) garantiza que se ubiquen suficientes instalaciones para cubrir la demanda, también contiene las tres restricciones (4),(5),(6) del modelo P-Mediana.

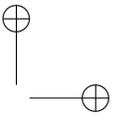
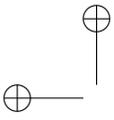
Se ha encontrado un estudio que aplica el problema de cobertura en la agroindustria, Marulanda *et al.* [40] buscan la ubicación de plantas de producción, su objetivo es la cobertura de una mayor demanda al menor costo, determinan el número necesario de plantas, con el fin de soportar la toma de decisiones para la expansión de una empresa. En las formulaciones no se ha tenido en cuenta la vida útil, los desperdicios, o el carácter perecedero de los productos agrícolas, el enfoque de las funciones ha sido la reducción de costos teniendo en cuenta la demanda, apertura y cierre de nuevas instalaciones. Sin embargo, el modelo propuesto por Marulanda [40] presenta un enfoque interesante teniendo en cuenta la capacidad limitada de las instalaciones a ubicar en la CS, generando una mayor aproximación al estado real de la misma, para lo cual agrega la ecuación (13).

$$\sum_j^m S_j Y_j \geq \sum_{i=1}^n d_i \quad (13)$$

### 3.1.3. Problemas P-Centro

El problema P-Centro es también conocido como MinMax, busca reducir al mínimo las distancias máximas entre cualquier demanda y su centro de distribución más cercano Este tipo de modelos





son utilizados para determinar la ubicación de una instalación sobre cualquier lugar de una red establecida, a los cuales se les denomina, problemas de centro absoluto [37]. Las soluciones al problema P-Centro se dividen típicamente en dos. La primera denominada centros de nodo en la que las instalaciones elegibles están restringidas a los vértices de la red y la segunda, centros absolutos en los que la instalación podría estar en cualquier lugar dentro de la red [26]. A los elementos que diferencia el modelo de los anteriores [31]:

$$\text{F.O:} \quad \text{Minimizar } Q \quad (14)$$

Restricciones:

$$\sum_{j \in J} d_{ij} y_{ij} - Q \leq 0 \quad \forall_j \in J, \quad (15)$$

La función objetivo junto a la restricción (14) minimiza la distancia máxima entre cliente y la instalación. Además de las restricciones (2), (3), (4), (5), (6), aparece la restricción (15) la cual limita el número de instalaciones a una cantidad preestablecida.

Los problemas P-Centro han sido trabajados tanto para productos agrícolas perecederos, para la ubicación de CD [41], como para la agroindustria en la localización de plantas de procesamiento encontrando el centro de gravedad del conjunto de clientes [5]. En las funciones objetivo del problema, en alimentos perecederos, se encuentran el minimizar los costos de ubicación y transporte [5] y la minimización del tiempo máximo del recorrido entre instalaciones [41]. Yang *et al.* [41] considera el carácter perecedero de los productos como una restricción adicional con tiempos de entrega. La duración de una entrega es importante para alimentos porque influyen en su calidad y la satisfacción del cliente, por lo tanto, los tiempos de viaje son considerados en el modelo como un factor de competencia. Para ello propone las siguientes ecuaciones en su modelo matemático, la (16) hace referencia a la función objetivo en donde se busca reducir el tiempo de transporte a lo largo de las rutas de la red y la (17) garantiza que el camino o ruta seleccionada es válida siempre y cuando exista asignación de nodos.

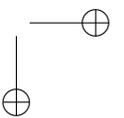
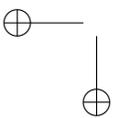
$$\text{Min} : C_r \{ (T_{ik} + dT_{km} + T_{mj}) Y_{ikmj} \leq \bar{Z} \} \quad (16)$$

$$Y_{ikmj} \geq Y_{ij} + Y_{jm} - 1 \quad (17)$$

Tanto Yang [41] como Boudahri [5] agregan al modelo decisiones de transporte con miras al mejoramiento del desempeño de la CS, adicionalmente Boudahri [5] incluye decisiones de asignación.

### 3.1.4. Problemas de localización Hub

Los problemas de localización Hub surgen en contextos en los que los clientes requieren el transporte entre un gran número de pares origen-destino, en las redes Hub se requiere que los clientes viajen desde su origen a un eje central y de allí se dirijan ya sea a su destino o a otro Hub y así sucesivamente hasta su destino final [26]. Las redes Hub consideran grafos no dirigidos y la existencia de flujo entre cada par de nodos. El número de Hubs puede ser fijo o variables. Si el conjunto de los centros se conoce, el problema tiene como solución la ruta más corta, de lo contrario el problema



es NP-complejo [35]. Klose [35] propone el modelo general. La función objetivo (18) minimiza los costos fijo y variables del envío y ubicación de las instalaciones.

La ecuación (19) asegura que la suma de las porciones de demanda no sea mayor al 100 %, la ecuación (20) controlan el número de instalaciones a ubicar.

$$\text{F.O:} \quad \min \sum_{i \in K} \sum_{k \in H} \sum_{m \in H} \sum_{j \in K} W_{ij} C_{ikmj} P_{ijp} + \sum_{j \in J} f_j x_j \quad (18)$$

Restricciones:

$$\sum_{i \in I} \sum_{j \in J} p_{ijp} = 1 \quad \forall p \in P \quad (19)$$

$$0 \leq p_{ijp} \leq x_j \quad \forall i, j, p \quad (20)$$

Para la logística de los productos perecederos es importante una buena ubicación de las instalaciones en centros de distribución urbana eficaces y viables [42]. Los problemas de localización tipo Hub en alimentos agrícolas perecederos buscan el diseño de una red de distribución al por mayor con conexiones eficientes. Se han trabajado diferentes tipos de objetivos como la minimización de los costos totales de la red incluyendo costos de transporte y localización, la minimización del tiempo máximo del recorrido entre instalaciones [43].

Para productos perecederos se han tenido en cuenta los tiempos y la distancia de envío, la satisfacción de la demanda total [43] y la duración en las entregas como factor de competencia al relacionarlo con la calidad y satisfacción del cliente, la distancia máxima de entrega, capacidad limitada de los Hub, consolidación del flujo que se envía y escenarios difusos en los tiempos de viaje. [41]. Etemadnia et al. [43] en su modelo considera la cantidad de producción y demanda (21) (22), y la capacidad de las instalaciones tanto con límite superior como inferior (23-24).

$$\sum_h m s_{ij} \leq P P_i \quad (21)$$

$$\sum_h m d_{jh} \leq d_j \quad (22)$$

$$x_h L_h \leq \sum_h m d_{jh} \leq x_h U_h \quad (23)$$

$$x_j L_j \leq \sum_h m s_{ij} \leq x_j U_j \quad (24)$$

### 3.1.5. Problema multi-producto

En aquellos casos donde no sea posible agregar la demanda, la producción, la manipulación y los costos de distribución, será necesario acudir a un modelo multi-producto donde, por ejemplo, la capacidad de los nodos, la demanda y los flujos están separados con respecto a grupos de productos homogéneos [35]. Klose [35] presenta el siguiente modelo general de este tipo de problemas:

$$\min \sum_{i \in I} \sum_{j \in J} \sum_{p \in P} q_{ij} p_{ijp} + \sum_{i \in I} \sum_{j \in J} g_{ij} y_{ij} + \sum_{i \in J} f_j x_j \quad (25)$$

$$y_{ij} - x_j \leq 0 \quad \forall i \in I, j \in J \quad (26)$$

$$p_{ijp} - y_{ij} \leq 0 \quad \forall i \in I, j \in J, \forall p \in P \quad (27)$$

$$p_{ijp} \geq 0 \quad (28)$$

La función objetivo (25) minimiza los costos fijos y variables de proveer los productos, incluye la ecuación (19) presentada anteriormente, la (26) y (27) prohíbe asignar productos a los depósitos cerrados y que se entregue cuando no está disponible en el almacén. Mientras la restricción (28) controla que sea asignada por lo menos alguna porción de demanda a cada instalación.

En los problemas multi-producto en perecederos la función objetivo común es la minimización de los costos totales que incluyen los costos de ubicación, producción, transporte, los costos de implementación [2] y costos de inversión [44], se han contemplado modo de transporte, inventario, capacidad de las instalaciones [44], tipo de alimento a producir en cada planta y cantidad a enviar entre las instalaciones [2], algunos modelos incluso abarcan la planificación completa de la CS, considerando la incertidumbre de la demanda [44]. Se han enfocado en la reducción de distancias y costos, e incluso se ha incluido la congestión vehicular y los diversos de actores de una misma cadena. No se han encontrado investigaciones que tengan en cuenta el carácter perecedero de los productos, ni el estado de los mismos durante el movimiento en cadena.

Zhao [2] propone un modelo matemático en el que tiene en cuenta productos frescos, semi-procesados y procesados. Para cada uno de ellos construye una serie de restricción que calculan la cantidad de productos disponibles para suplir la demanda de los eslabones que intervienen. Dadas las características de este trabajo se expone a continuación el conjunto de restricciones que presentan mayor perecibilidad.

$$\sum_j \sum_n TPI_{ijn} \leq ppt * pp_i, \quad \forall i \in I \quad (29)$$

$$\sum_i \sum_n TPI_{ijn} \leq bigpos * x_j, \quad \forall j \in J \quad (30)$$

$$pp_i \leq ppm_i, \quad \forall i \in I \quad (31)$$

### 3.1.6. Problemas de servicios múltiples

En este tipo de problemas se distingue entre los modelos de elección del cliente y los modelos de asignación. Los de elección del cliente deciden la instalación donde ir, es un sistema apropiado para el sector minorista, en especial cuando el planificador de instalaciones también se encarga de decidir qué instalación suministra a un cliente determinado [26].

Neungmatcha *et al.* [45] proponen un modelo para la industria azucarera y la ubicación de estaciones teniendo en cuenta la planeación del transporte y distribución, el objetivo es la minimización de los costos de transporte y de operación, busca el número de instalaciones a abrir y relaciona a los cultivadores con las estaciones que cada uno de estos deberá suplir. En servicios múltiples los trabajos en CSA no se ha tenido en cuenta el factor percedero a lo largo del viaje del producto. El modelo se presenta a continuación:

$$f(x, y, z) = \sum_{i=1}^n \sum_{j=1}^m h d_{ij} Y_{ij} + \sum_{j=1}^m h l_{jM} y_j + \sum_{j=1}^m e_j y_j + \sum_{j=1}^m \sum_{c=1}^p o_c Z_{jc} + \sum_{j=1}^m \sum_{c=1}^p f_c Z_{jc} \quad (32)$$

$$\sum_{i=1}^n P P_i Y_{ij} \leq B_j y_j, \quad j = 1, 2, \dots, m \quad (33)$$

$$\sum_{j=1}^m y_j \leq M, \quad (34)$$

$$\sum_{i=1}^n P P_i Y_{ij} \leq \sum_{c=1}^p b_c Z_{jc}, \quad j = 1, 2, \dots, m \quad (35)$$

$$Y_{ij} = 0, \quad y_j = 0, \quad Z_{jc} \geq 0 \quad (36)$$

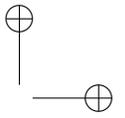
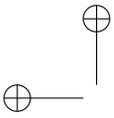
Este modelo se desarrolló con el objetivo de minimizar el costo total en cualquier período en un año agrícola (32), incluye el costo de transporte de un campo a una estación de carga y desde la estación de carga de un ingenio azucarero. La restricción (33) asegura que la capacidad de servicio de cada estación no deberá exceder su capacidad y utiliza la ecuación la (2) que asegura que todos los campos se asignaran a una sola estación, la (34) restringe el número total de estaciones abiertas, para que no exceda el número total de estaciones candidatas, la (35) asegura que la cantidad total de la caña de azúcar se puedan atender por las estaciones.

### 3.2. Tipos de modelado y métodos de solución

Una vez identificados los problemas de localización aplicados a alimentos se procedió a revisar los diferentes tipos de modelado matemático y los métodos de solución utilizados con énfasis en perecederos. Los algoritmos de solución incluyen, entre otros, la búsqueda local y enfoques basados en programación matemática. [35]. En la revisión literaria se identificaron que los autores han utilizado para el problema de alimentos agrícolas perecederos modelos de programación lineal, programación entera mixta (lineal y no lineal) y programación entera binaria y, los métodos utilizados para la solución algoritmos genéticos, heurísticas y meta-heurísticas.

#### 3.2.1. Programación lineal

Boudahri *et al.* [5] presentan un modelo de programación lineal que busca encontrar el centro de gravedad de los grupos de clientes, ubicar y asignar los mataderos en la cadena de distribución de carne de pollo, tiene en cuenta la capacidad de las instalaciones y de los vehículos. El objetivo es



minimizar los costos de ubicación y de transporte. El modelo es capacitado y contempla las coordenadas geográficas de cada uno de los puntos posibles en seleccionar y asignar, la distancia entre nodos y el transporte.

Tong *et al.* [11] proponen un modelo de localización que tiene como objetivo seleccionar los lugares y los horarios de servicios de mercados de agricultores realizados directamente al cliente final de la CS. El modelo propuesto tiene en cuenta la variabilidad de la demanda a lo largo del tiempo, la distancia y el número de viajes de los clientes hacia dichos lugares de venta. El problema es no capacitado y trata aspectos como la red de distribución, la distancia entre nodos y el crecimiento poblacional.

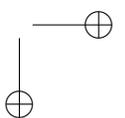
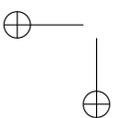
### 3.2.2. Programación entera mixta

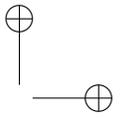
Hiassat y Diabat [46] plantean un modelo para la ubicación de las instalaciones que afectan a la empresa durante períodos de tiempo significativamente más largos que otros. Determinan cuantos almacenes hay que abrir, donde localizarlos y que clientes asignar a los mismos. Existe un solo proveedor que distribuye a múltiples minoristas con demanda determinística a través de un conjunto de centros de distribución, de un producto perecedero con cierta vida útil. El modelo incluye la localización, los inventarios y el costo de ruteo. El objetivo es reducir el costo.

Zhao y Lv [47] diseñan un modelo para la ubicación de instalaciones, en donde se selecciona la capacidad de producción y el modo de transporte para el diseño de la CSA de manzanas. El objetivo es reducir los costos de producción y transporte. La solución se realiza a través de enjambre de partículas. Zhao y Dou [2] proponen un modelo para optimizar la CSA agrícolas con el objetivo de reducir los costos de producción y transporte de la cadena. Con el modelo buscan ubicar las instalaciones de manera óptima y con ello determinar la capacidad máxima de producción, el tipo de alimento a producir en cada planta, la selección de la cantidad a transportar y el modo de transporte utilizado en la distribución. Para la solución del problema se utiliza enjambre de partículas. Yang *et al.* [41] presentan un modelo de programación entera mixta p-hub en entornos difusos (suponen elementos difusos en el tiempo de viaje y entre los nodos de origen-destino). El objetivo es minimizar el tiempo de recorrido máximo entre cada instalación. Se resuelve a través de enjambre de partículas.

Neungmatcha *et al.* [45] presentan un modelo solucionado a través de un algoritmo genético adaptativo, en el cual tienen como principal objetivo hallar la ubicación óptima para la instalación de las estaciones en una CS de la caña de azúcar, intentando maximizar el uso de la capacidad total. El modelo es tratado como un problema multi- servicio donde se tiene en cuenta el tipo y número de estos, y se espera minimizar el costo total. El modelo considera la capacidad de las instalaciones y la asignación de los clientes.

Govindan *et al.* [4] proponen un modelo de optimización multi-objetivo en una CSA de perecederos, el cual integra la sostenibilidad, el cálculo del número y ubicación de las instalaciones, optimiza la cantidad de los productos entregados y las rutas a través de las cuales se realiza la distribución. Los autores proponen un modelo capacitado de doble eslabón con ventanas de tiempo, buscando reducir las emisiones de gases de efecto invernadero, utiliza para la solución un híbrido





de dos algoritmos (Multi objective particle swarm optimization - Multi objective variable neighborhood search).

Etemadnia *et al.* [43] examinan la ubicación de instalaciones en una CSA perecederos para facilitar la transferencia entre los lugares de producción y consumo. Para ello proponen un modelo de programación entera mixta donde el objetivo es la minimización del costo total de la red, incluyendo los gastos de transporte y el costo de localizar cada una de las instalaciones. Adicional a esto los autores realizan un estudio por escenarios variando parámetros tales como, las distancias de viaje y la capacidad de cada instalación. La solución del problema se hace a través de una heurística.

### 3.2.3. Programación entera mixta no lineal

Zhi-lin y Dong [38] desarrollan un modelo de optimización para la ubicación de instalaciones. Se centran en el problema de seleccionar la mejor ubicación para el centro de distribución de alimentos agrícolas en China, que represente el menor costo posible, otorgando un análisis económico en cuanto a la construcción y operación de dicho centro. A diferencia de otros los trabajos en programación lineal, los autores tienen en cuenta una tasa de descomposición constante, sin embargo, no relaciona un costo de pérdida. Inicialmente se conocen los puntos de ubicación y no tiene en cuenta la capacidad de las instalaciones.

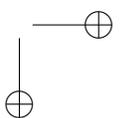
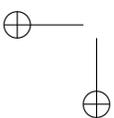
Jouzani *et al.* [44] proponen un modelo para la ubicación de instalaciones y la planificación de la CS de leche. El objetivo del modelo es la ubicación óptima de instalaciones y determinan los volúmenes de producción con bajos costos de transporte, consideran el tráfico de las vías y la incertidumbre de la demanda. Tienen en cuenta los posibles cambios en el sistema de transporte y producción, la cantidad de alimento a procesar y el inventario a mantener. Los puntos de ubicación se establecen en una red conocida.

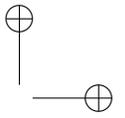
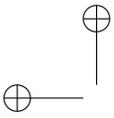
Firoozi *et al.* [10] trabajan un problema de integración entre el control de inventarios y la ubicación de instalaciones. Los autores consideran de importancia aspectos de manera simultánea en la toma de decisiones, consideran en el costo de transporte y reabastecimiento, afectando de esta manera la política óptima de inventario. Para la solución del problema se hace uso de un algoritmo memético. Se considera que las demandas en los centros de distribución son estocásticas, se manejan puntos potenciales de ubicación conocidos y costos de transporte.

Damghani *et al.* [48] consideran que la ubicación de almacenes y el enrutamiento de vehículos son cuestiones esenciales para la distribución de productos perecederos. Proponen un modelo de programación entera mixta bi-objetivo. Donde los objetivos a trabajar son reducir el costo total de la CS y equilibrar la carga de trabajo en los centros de distribución. Para resolver el problema se usa un algoritmo genético tipo II (NSGA-II). El problema tiene en cuenta la capacidad de los centros de distribución, el transporte y la eficiencia en los resultados obtenidos en la cadena.

### 3.2.4. Programación entera binaria

Weimin *et al.* [9] consideran un problema de diseño de red de distribución donde se busca decidir los sitios en los cuales deben estar ubicados los centros de distribución de productos agrícolas pe-





recederos. Con la premisa de que cada centro tiene la capacidad suficiente para servir a su zona de consumo, además de que cada zona acepta productos de solo un centro de distribución. El objetivo es minimizar el costo total, el cual incluye los costos de ubicación, los costos de inventario, los costos de transporte y los costos de pérdida por deterioro. Para la solución del problema utilizan un algoritmo genético.

Como se observa algunos de los trabajos reseñados no solamente describen con claridad el tipo de modelado matemático utilizado, también mencionan y representan la forma y métodos utilizados en la solución y optimización de los problemas. Se identificó que, para la solución de los problemas de localización de instalaciones en la cadena agrícola de perecederos, se utilizan en general: algoritmos genéticos, heurísticas y meta-heurísticas.

A continuación se realiza una clasificación con aquellos trabajos en los cuales no se presenta con claridad el tipo de modelado matemático, pero si la forma de solución.

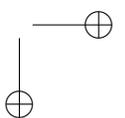
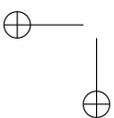
### 3.2.5. Heurísticas y meta-heurísticas

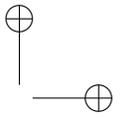
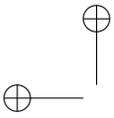
Federgruen *et al.* [49] presentan un modelo de asignación de productos perecederos, distribuido desde un centro regional para un conjunto dado de ubicaciones con demanda aleatoria. Consideran un problema combinado de asignación de inventario disponible en dicho centro y la forma como se van a realizar las entregas. Se utiliza un algoritmo para la solución y se busca la reducción del costo en cuanto a distribución y transporte.

Zhang y Yang [7] plantean un modelo de optimización para la ubicación de instalaciones en el sistema de emergencia de productos perecederos básicos. Dicho modelo tiene como objetivo la minimización de los costos logísticos bajo limitaciones de tiempo, consideran los costos de transporte y aquellos que representan la pérdida de producto a lo largo de la red de distribución. Los autores establecen un centro de servicio en un área determinada con el fin de asegurar que se abastece la demanda, es decir un punto de demanda es atendido por una instalación, las instalaciones se pueden ubicar tanto en los bordes de la red, como en los vértices, con el fin de lograr la eficiencia de toda la red. Se plantea una heurística para la solución del problema.

Xiaohui y Wen [39] propusieron un modelo para la selección del lugar más adecuado para ubicar los centros de distribución tanto para frutas y verduras en Beijing. El modelo considera la relación espacio-tiempo, con el objetivo de eliminar la discrepancia en el tiempo y lugar entre el productor y el consumidor final del alimento, reducen los costos logísticos y los desperdicios, ofreciendo de manera conveniente y eficiente las frutas a los diversos mercados. En este trabajo se utilizó una heurística con el propósito de reducir la complejidad del problema.

Huang y Xie [50] proponen un modelo de solución al problema de ubicación de centros de distribución en la logística alimentaria de emergencia. En la decisión de localización se tomó en cuenta la eficiencia, la seguridad, la fiabilidad y la incertidumbre en los eventos de emergencia. Desarrollan el modelo bajo la premisa de que la correcta ubicación de los centros de distribución es importante para asegurar que el sistema tenga alta eficiencia, lo que significa una respuesta más rápida, menor tiempo de entrega y costo de operación más bajo cuando ocurren eventos de emergencia. Para la





solución del problema se combinan redes neuronales con un algoritmo de colonia de hormigas, dentro de una red de posibles nodos a seleccionar.

Marulanda *et al.* [40] trabajan el problema de localización de plantas de producción con restricción de capacidad, lo aplican a una empresa que produce y comercializa productos de panadería y repostería en diversas regiones de Colombia. Buscan cubrir la mayor demanda posible con la instalación de una o más plantas de producción al menor costo. El problema es solucionado a través de Relajación Lagrangeana y Búsqueda Tabú.

Diatha *et al.* [51] proponen un modelo de formulación matemática multi-objetivo de ruteo y localización para en el mercado de hongos, lo solucionan con una heurística. Se conocen los posibles lugares donde se ubicarán los centros de distribución y se asigna un canal de distribución. Se busca minimizar los gastos, las distancias y los tiempos de respuesta, teniendo en cuenta ventanas tiempo.

Drezner y Scott [52] plantean un modelo que combina decisiones de inventario y localización de instalaciones para productos perecederos. Consideran la ubicación de un único centro de distribución que sirve un número finito de puntos de venta. El objetivo del modelo es reducir al mínimo el costo total, compuesto por los costos de inventario y los costos de transporte desde el centro de distribución hasta los puntos de venta. Se conocen los puntos potenciales para la ubicación de las instalaciones, la demanda se modela con una función de densidad y se utilizan las Heurística GWA (Algoritmo Weiszfeld Generalizado) y BTST (triángulo grande, triángulo pequeño), para la solución.

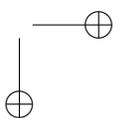
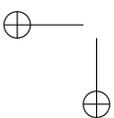
### 3.2.6. Algoritmos genéticos

Gong *et al.* [53] analizan de manera conjunta el problema de inventarios e instalación de ubicaciones para productos perecederos agrícolas en China. A través de un algoritmo, buscan reducir al mínimo el costo total de transporte, el nivel de inventarios y la cantidad de desperdicios a lo largo de la CS, encuentran la mejor distribución de los productos.

Chen y Zhong [15] describen un algoritmo genético mejorado para la solución de problemas de ubicación de centros de distribución de productos perecederos, consideran áreas restrictivas, tales como lagos y zonas administrativas. Los autores incluyen el impacto que tiene el carácter perecedero de los productos sobre el costo de distribución global. En el modelo se buscan establecer las zonas de distribución y sus respectivas rutas de manera factible.

## 4. Discusión

A lo largo de la revisión se ha destacado que los productos agrícolas tienen un alto nivel de consumo, un proceso de distribución complejo, un precio bajo, una vida útil corta y difícil gestión de almacén, lo cual según [9] se ve reflejado en mayores costos logísticos. Sin embargo, la cantidad de publicaciones que se encuentran sobre el tema es escasa, comparada con la literatura referente a cadenas no agrícolas. Esto podría deberse a la complejidad que representa el modelado en este tipo de cadenas dado el número de eslabones y actores que participan, así como las variables es-



pecíficas a tener en cuenta. Como se evidencia a lo largo del artículo de revisión aspectos como la seguridad alimentaria, impacto ambiental y el factor perecedero de los productos no se ha tratado ampliamente, a pesar de que tal como lo plantea [32] en los problemas de localización son aspectos críticos para garantizar que una CSA tenga altas medidas de desempeño. En la Tabla II, en la cual se presentan los principales aspectos tenidos en cuenta por cada uno de los autores en los artículos de productos perecederos consultados.

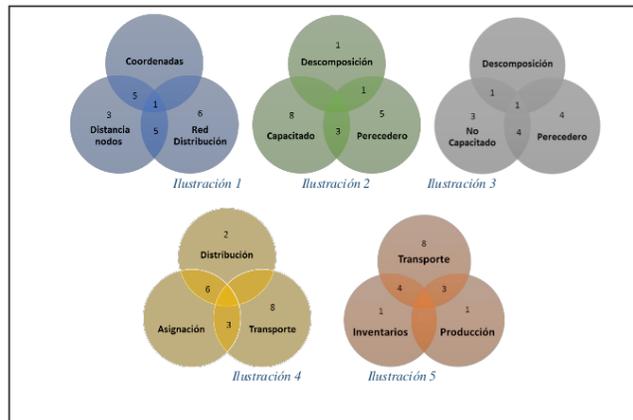
**Tabla II.** Aspectos Tratados en los Artículos

Referencia	Año	Aspectos Tratados en los artículos																					
		Coor	R.D.	N Ca	Ca	D.N	D	Des	Per	T	S.A	Efi	Fia	C	Inv	V.D	Imp A	T.T	P	Pob	Asig	E.D	
Gong	2007				x	x			x	x					x	x							
Zhin	2007	x		x		x	x	x															
Zhang	2007		x	x		x			x														
Xiaohui	2009	x	x	x		x			x	x													
Huang	2009		x	x							x	x	x										
Marulanda	2010		x		x				x											x			
Hiassat	2010		x						x	x					x	x						x	
Weimin	2011			x				x	x	x				x									
Zhao	2011				x		x			x										x			
Boudahri	2011	x			x	x				x												x	
Zhao	2011		x			x	x			x										x			
Yang	2012		x	x		x			x	x													x
Tong	2012		x	x		x															x		
Diatha	2012				x	x	x			x													
Drezner	2013	x		x		x	x							x	x								
Chen	2013	x		x		x			x	x				x									
Jouzdani	2013		x		x		x			x		x			x					x			
Neungmatcha	2013		x		x	x				x													x
Firoozi	2014		x							x				x	x								
Godivan	2014				x	x				x				x		x	x						
Morganti	2014		x		x		x		x	x													
Damghani	2015				x					x		x											
Etemadnia	2015	x			x	x	x			x													
Total		6	12	9	11	14	8	2	9	17	1	3	1	2	5	5	1	1	4	1	3	1	

Coor: Coordenadas; R.D: Red de Distribución; NCa: No capacitado; Ca: Capacitado; D.N: Distancia nodos; D: Distribución; Des: Descomposición; Per: Perecedero; T: Transporte; S.A: Seguridad Alimentaria; Efi: Eficiencia; Fia: Fiabilidad; C: Calidad; Inv: Inventario; V.D: Variabilidad Demanda; ImpA: Impacto Ambiental; T.T: Tiempo de trabajo; P: Producción; Pob: Población; Asig: Asignación; E.D: Entornos Difusos.

A partir de la revisión de la literatura se propone una tipología mediante de las variables tratadas en los artículos según tipos y cantidad empleada las cuales son presentadas en la figura 2. En la Ilustración 1, se muestra la forma en que los autores conciben la relación de distancia entre nodos candidatos e instalaciones, las variables utilizadas en este tipo son coordenadas, distancia entre nodos y red de distribución. En la Ilustración 2 se encuentran las variables específicas de los productos perecederos, mostrando el número de artículos que han tenido en cuenta el carácter de perecibilidad, la descomposición del producto y la capacidad de las instalaciones.

Por otra parte, la Ilustración 3 clasifican en modelos no capacitados, se observa que existe un mayor número de artículos que tienen en cuenta la capacidad de las instalaciones, pero en pocas ocasiones se considera la perecibilidad y la descomposición. Las Ilustraciones 4 y 5 clasifican el problema de localización según los “drivers” logísticos, si se han tratado de manera individual o como diadas, la Ilustración 4 contempla las decisiones de distribución, asignación, mientras la 5 las



**Figura 2.** Principales variables trabajadas en los artículos.

referidas a transporte o ruteo, producción e inventarios, siendo transporte el de mayor frecuencia.

Finalizada la definición de la tipología según las variables que se utilizan con mayor frecuencia en la literatura científica de los problemas de localización de productos perecederos, se realiza una clasificación de los artículos de acuerdo al tipo de problema, modelo matemático y tipo solución que se han utilizado Tabla III.

Se identifica que el tipo de modelo matemático más utilizado es la programación entera mixta, esto se debe a un número grande de variables, parámetros y subíndices que los modelos generalmente incluyen, al contemplar características propias de las CS de productos agrícolas de perecederos. De igual manera se encuentra que dado la complejidad de los modelos, NP-Hard, se hace necesario la utilización de heurísticas y Meta-heurísticas que aproximen una solución al óptimo de cada modelo planteado.

En lo que respecta a las funciones objetivos, los modelos utilizan principalmente la minimización del costo total, incluyen costos de transporte, inventario y de producción. No es común la inclusión, en la formulación matemática, del carácter perecedero de los alimentos, en algunos casos se representa dicha característica como un costo, un tiempo máximo de transporte ligado a la vida útil o como una función de deterioro del mismo

## 5. Conclusiones

El interés en las cadenas agroalimentarias se ha incrementado principalmente por la importancia que ha adquirido el consumo de productos frescos alrededor del mundo. Decisiones estratégicas relevantes como la localización de instalaciones son de importancia durante el diseño de la CS y de su red de distribución dado el carácter perecedero de los productos frescos. Sin embargo, el tema ha sido poco tratado debido especialmente a la dificultad para el modelado, así como el tipo y cantidad de variables a considerar en este tipo de CS.

**Tabla III.** Taxonomía de clasificación de los artículos de productos perecederos

Referencia	Año	Tipos de Problemas					Tipos de Modelos Matemáticos				Tipos de Soluciones		
		P-Mediana	P-Cobertura	P-Centro	Hubs	Multiproducto	Servicios Múltiples	Programación lineal	Programación entera mixta	Programación entera mixta no lineal	Programación binaria	Heurísticas y Metaheurísticas	Algoritmos Genéticos
Gong	2007											x	
Zhin	2007	x								x			
Zhang	2007	x									x		
Xiaohui	2009	x									x		
Huang	2009										x		
Marulanda	2010		x								x		
Hassat	2010								x				
Weimin	2011									x			
Zhao	2011								x				
Boudahri	2011			x				x					
Zhao	2011					x			x				
Yang	2012			x	x				x				
Tong	2012	x						x					
Diatla	2012										x		
Drezner	2013										x		
Chen	2013											x	
Jouzdani	2013					x				x			
Neungmatcha	2013						x		x				
Firozi	2014									x			
Godivan	2014								x				
Morganti	2014				x								
Damghani	2015									x			
Etamadnia	2015				x				x				
Total		4	1	2	3	2	1	2	7	4	1	6	2

Todo esto conlleva que a pesar de que en los últimos años ha ido en aumento el número de investigaciones y trabajos sobre las cadenas agroalimentarias a lo largo del mundo, principalmente en países desarrollados que buscan fortalecerse especialmente en frutas y verduras, aún existe espacio para investigaciones y propuestas en donde se incluyan en un solo modelo variables como capacidad, tipo de instalación, percibibilidad de los alimentos, pérdidas y comportamiento estocástico de la demanda, con el fin de buscar modelos que permitan la toma de decisiones para mejorar el desempeño de las cadenas.

En estos países es necesaria la correcta ubicación de los centros de acopio, centrales mayoristas, empresas procesadoras, las plataformas logísticas y los puertos, así como la ubicación de los consumidores, en los modelos deben contemplarse su funcionalidad, evolución de capacidad y un adecuado diseño de modelos de gestión, contemplando las costumbres comerciales y sus implicaciones en los flujos, tales como la infraestructura y características de la flota. Existen alimentos perecederos como las frutas, en los cuales no se ha tenido en cuenta aspectos como los cambios de temperatura, humedad relativa y pisos térmicos, esenciales para garantizar la vida útil de la fruta y su calidad, que en países con cordilleras como Colombia se hace fundamental. Esto abre puertas a futuras investigaciones que modelen el carácter perecedero de las frutas teniendo en cuenta cambios climáticos, factores internos y externos, no solo al penalizarlo como un costo adicional sino que realmente represente los estados y las condiciones en su viaje a través de los diferentes eslabones de la CSA.

Es importante que en los modelos futuros se contemple tipos de transporte que permitan conservar durante el mayor tiempo posible las características organolépticas de los alimentos, tales como atmósferas modificadas y controladas, actualmente se limitan a identificar el tiempo en el cual la fruta pierde su calidad. Analizando la relación costo-beneficio al reducir las pérdidas de producto por descomposición y aumentando la calidad del producto. La mayor parte de la pérdida después de la cosecha ocurre en los procesos de almacenamiento, empaquetamiento y distribución por debilidades en la provisión y coordinación del servicio logístico.

En este estudio se logró identificar los principales modelos, tipologías de modelado y métodos de solución enfocados a las cadenas agrícolas de productos perecederos, exponiendo oportunidades de mejora y brechas en el conocimiento en la investigación científica de este tipo de problemas, con la intención de que puedan ser tratados en proyectos futuros principalmente en países montañosos como Colombia [55], [56] que permitan responder a la demanda de productos frescos conservando la calidad de estos. Asimismo, se identifica que los modelos dinámicos no son empleados para decisiones sobre localización de instalaciones en las CSA de perecederos, no obstante las especificidades de las mismas a través del tiempo. Por lo tanto, este es otro campo abierto para los investigadores.

## Referencias

- [1] J. A. Orjuela C, C. A. Castañeda C y M. E. Calderón, "Análisis de la Cadena de valor en las estructuras productivas de uchuva y tomate de árbol en la provincia de Sumapaz y el Distrito Capital". *Ingeniería*, vol. 13, no 2, pp. 4-12, 2008. ↑24
- [2] X. Zhao y J. Dou, "A Hybrid Particle Swarm Optimization Approach for Design of Agri-Food Supply Chain Network". *Proceedings of 2011 IEEE International Conference on Service Operations, Logistics and Informatics*, pp. 162-167, 2011. ↑24, 25, 29, 33, 35
- [3] J. A. Orjuela Castro y I. A. Chavarrio Colmenares, "Caracterización de la cadena de abastecimiento de panela para la provincia de Bajo Magdalena-Cundinamarca". *Ingeniería*, vol. 16, no 2, pp. 107-124, 2011. ↑24
- [4] K. Govindan, A. Jafarian, R. Khodaverdi y K. Devika, "Two-Echelon Multiple-Vehicle Location-Routing Problem with Time Windows for Optimization of Sustainable Supply Chain Network of Perishable Food". *International Journal of Production Economics*, pp. 9-28, 2014. ↑24, 25, 35
- [5] F. Boudahri, Z. Sari, F. Maliki y M. Bennekrouf, "Design and Optimization of the Supply Chain of Agri-Foods: Application Distribution Network of Chicken Meat". *2011 International Conference on Communications, Computing and Control Applications, CCCA 2011*, 2011. ↑24, 31, 34
- [6] M. M. Herrera Ramírez y J. A. Orjuela Castro, "Perspectiva de Trazabilidad en la cadena de suministros de frutas: un enfoque desde la dinámica de sistemas". *Ingeniería*, vol. 19, no 2, 2014. ↑24
- [7] M. Zhang y J. Yang, "Optimization Modeling and Algorithm of Facility Location Problem in Perishable Commodities Emergency System". *Proceedings - Third International Conference on Natural Computation*, pp. 246-250, 2007. ↑24, 25, 29, 37
- [8] O. Ahumada y J. R. Villalobos, "Application of Planning Models in the Agri-Food Supply Chain: a Review". *European Journal of Operational Research*, p. 5, 2009. ↑24
- [9] F. A. Pérez Mantilla y F. J. Torres Delgado, "Modelos de inventarios con productos perecederos: revisión de literatura". *Ingeniería*, vol. 19, no 2, 2014. ↑24, 25, 36, 38
- [10] W. Di, J. Wang, B. Li y M. Wang, "A Location-Inventory Model for Perishable Agricultural Product Distribution Centers". *2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce, AIMSEC 2011 - Proceedings*, pp. 919-922, 2011. ↑25, 36
- [11] A. F. Ruiz Moreno, A. L. Caicedo Otavo y J. A. Orjuela Castro, "Integración externa en las cadenas de suministro agroindustriales: una revisión al estado del arte". *Ingeniería*, vol. 20, no 2, pp. 9-30, 2015. ↑25, 29, 30, 35
- [12] Z. Firoozi, N. Ismail, S. Ariaifar, S. H. Tang, M. K. M. A. Ariffin y A. Memariani, "Effects of Integration on the Cost Reduction in Distribution Network Design for Perishable Products". *Mathematical Problems in Engineering*, pp. 1- 10, 2014. ↑25
- [13] D. Tong, F. Ren y J. Mack, "'Locating Farmers' Markets with an Incorporation of Spatio-Temporal Variation". *Socio-Economic Planning Sciences*, pp. 149-156, 2012. ↑25
- [14] M. T. Lucas y D. Chhajed, "Applications of Location Analysis in Agriculture: a Survey". *Journal of the Operational Research Society*, pp. 561-578, 2004. ↑25
- [15] S. Validi, A. Bhattacharya y P. J. Byrne, "A Case Analysis of a Sustainable Food Supply Chain Distribution System-a Multi-Objective Approach". *International Journal of Production Economics*, pp. 71-87, 2014. ↑25, 38
- [16] R. Akkerman, P. Farahani y M. Grunow, "Quality, Safety and Sustainability in Food Distribution: a Review of Quantitative Operations Management Approaches and Challenges". *OR Spectrum*, pp. 863-904, 2010. ↑25

- [17] X. Chen y C.-y. Zhong, "An Improved Genetic Algorithm for Location Problem of Logistic Distribution Center for Perishable Products". *International Asia Conference on Industrial Engineering and Management Innovation (IEMI2012) Proceedings*, pp. 949-959, 2013. ↑25
- [18] S. Chopra y P. Meindl, *Administración de la cadena de suministro*, México, PEARSON Prentice Hall, 2008, p. 554. ↑25
- [19] B. M. Beamon, "Supply Chain Design and Analysis: Models and Methods". *International Journal of Production Economics*, pp. 281-294, 1998. ↑25
- [20] Z. Ling y M. Zhang, "Optimization of Supply Chain Design Based on Knowledge Discovery of Distribution Network". *2009 International Joint Conference on Artificial Intelligence*, pp. 769-772, 2009. ↑25
- [21] S. Yu, "Supply Chain Distribution Network - Optimized Design Model and Its Solution Based on Time Constraints". *2012 Second International Conference on Intelligent System Design and Engineering Application*, pp. 941-945, 2012. ↑25
- [22] I. J. Moreno Castañeda, J. L. Roa Canal y J. A. Orjuela Castro, "Caracterización de la logística de la cadena de abastecimiento agroindustrial frutícola en Colombia". Bogotá, D. C., 2012. ↑25
- [23] MinSalud; FAO, *Perfil nacional de consumo de frutas y verduras*. Autor, 2013. ↑26
- [24] CORPOICA, Escenario actual de CTI en el sector hortofrutícola : análisis y retos futuros. Autor, 2013. ↑27
- [25] H.-S. Hwang, "A Food Distribution Model for Famine Relief". *Computers & Industrial Engineering*, pp. 335-338, 1999. ↑27
- [26] R. Carro Paz y D. González Gómez, *Localización de instalaciones*. Buenos Aires, 2012. ↑27, 28, 30, 31, 33
- [27] Ş. Erengüç, N. Simpson y A. J. Vakharia, "Integrated Production/Distribution Planning in Supply Chains: an Invited Review". *European Journal of Operational Research*, pp. 219-236, 1999. ↑27
- [28] C. E. H. ReVelle, "Location Analysis: a Synthesis and Survey". *European Journal of Operational Research*, pp. 1-19, 2005. ↑27
- [29] A. Weber, "Über den Standort der Industrien. I. Teil: Reine Theorie des Standortes, Tu bingen, Germany, Translated as On the Location of Industries". p. 153, 1909. ↑27
- [30] H. Hotelling, "Stability in Competition". *The Economic Journal*, pp. 41-57, 1929. ↑27
- [31] W. Christaller, *Central Places in Southern Germany*. Prentice-Hall, 1933. ↑27, 28, 30, 31
- [32] A. Losch, *The Economics of Location*. New Haven, Yale University Press, 1954. ↑27, 39
- [33] C. ReVelle, H. Eiselt y M. Daskin, "A Bibliography For Some Fundamental Problem Categories In Discrete Location Science". *European Journal of Operational Research*, pp. 817-848, 2008. ↑27, 28
- [34] Q. Meng, Y. Huang y R. L. Cheu, "Competitive Facility Location on Decentralized Supply Chains". *European Journal of Operational Research*, pp. 487-499, 2009. ↑27
- [35] M. Melo, S. Nickel y F. Saldanha-da-Gama, "Facility Location And Supply Chain Management—A Review". *European Journal of Operational Research*, pp. 401-412, 2009. ↑27, 28, 32, 34
- [36] M. Melo, S. Nickel y F. Saldanha da Gama, "Dynamic Multi-Commodity Capacitated Facility Location: a Mathematical Modeling Framework for Strategic Supply Chain Planning". *Computers & Operations Research*, pp. 181-208, 2006. ↑28
- [37] A. Klose y A. Drexl, "Facility Location Models for Distribution System Design". *European Journal of Operational Research*, pp. 4-29, 2005. ↑28, 31
- [38] R. Sridharan, "The Capacitated Plant Location Problem". *European Journal of Operational Research*, pp. 203-213, 1995. ↑29, 36
- [39] S. H. Owen y M. S. Daskin, "Strategic Facility Location: a Review". *European Journal of Operational Research*, pp. 423-447, 1998. ↑29, 37
- [40] S. Z.-I. S. Zhi-lin y W. D. W. Dong, "Location Model of Agricultural Product Distribution Center". *2007 International Conference on Management Science and Engineering*, pp. 117-120, 2007. ↑30, 38
- [41] Q. Xiaohui y Y. Wen, "Studies on Spatio-Temporal Collaboration Model for Location Analysis of Vegetable & Fruit Logistics". *6th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2009*, pp. 619-626, 2009. ↑31, 32, 35
- [42] M. V. Marulanda, G. I. Leguizamón y K. Y. Niño Mora, "Solución al problema de localización (cflp) a través de búsqueda tabú y relajación lagrangeana, caso de estudio: industria de productos alimentarios". *Purntr*, 2010. ↑32
- [43] K. Yang, Y. Liu y G. Yang, "An Improved Hybrid Particle Swarm Optimization Algorithm for Fuzzy P-Hub Center Problem". *Computers & Industrial Engineering*, pp. 133-142, 2012. ↑32, 36
- [44] E. Morganti y J. Gonzalez-Feliu, "City Logistics For Perishable Products. The Case of the Parma's Food Hub". *Case Studies on Transport Policy*, 2014. ↑33, 36

- [45] H. Etemadnia, S. J. Goetz, P. Canning y M. S. Tavallali, "Optimal Wholesale Facilities Location within the Fruit and Vegetables Supply Chain With Bimodal Transportation Options: an LP-MIP Heuristic Approach". *European Journal of Operational Research*, pp. 648-661, 2015. ↑34, 35
- [46] J. Jouzdani, S. J. Sadjadi y M. Fathian, "Dynamic Dairy Facility Location and Supply Chain Planning Under Traffic Congestion and Demand Uncertainty: a Case Study of Tehran". *Applied Mathematical Modelling*, pp. 8467-8483, 2013. ↑35
- [47] W. Neungmatcha, K. Sethanan, M. Gen y S. Theerakulpisut, "Adaptive Genetic Algorithm for Solving Sugarcane Loading Stations with Multi-Facility Services Problem". *Computers and Electronics in Agriculture*, pp. 85-99, 2013. ↑35
- [48] A. H. Hiassat y A. Diabat, "A Location Inventory Routing Problem with Perishable Products". *Proceedings of the 41st International Conference on Computers & Industrial Engineering*, pp. 386-391, 2010. ↑36
- [49] X. Zhao y Q. Lv, "Optimal Design of Agri-Food Chain Network: An Improved Particle Swarm Optimization Approach". *2011 International Conference on Management and Service Science*, pp. 1-5, 2011. ↑37
- [50] K. Khalili-Damghani, A.-R. Abtahi y A. Ghasemi, "A New Bi-objective Location-routing Problem for Distribution of Perishable Products: Evolutionary Computation Approach". *Journal of Mathematical Modelling and Algorithms in Operations Research*, 2015. ↑37
- [51] A. Federgruen, G. Prastacos y P. H. Zipkin, "An Allocation and Distribution Model for Perishable Products". *Operations Research*, pp. 75-82, 1986. ↑38
- [52] X. R. Huang y R. H. Xie, "A Model on Location Decision for Distribution Centers of Emergency Food Logistics". *2009 2nd International Conference on Information and Computing Science, ICIC 2009*, pp. 232-235, 2009. ↑38
- [53] K. Diatha, R. Karumanchi y S. Garg, "Mobile Enabled Operations Management Using Multi-Objective Based Logistics Planning For Perishable Products". *Computers and Industrial Engineering* 42, pp. 16-18, 2012. ↑38
- [54] Z. Drezner y C. H. Scott, "Location of a distribution center for a perishable product". *Mathematical Methods of Operations Research*, pp. 301-314, 2013. ↑
- [55] W. Gong, D. Li, X. Liu, J. Yue y Z. Fu, "Improved Two-Grade Delayed Particle Swarm Optimisation (TGD-PSO) for Inventory Facility Location for Perishable Food Distribution Centres in Beijing". *New Zealand Journal of Agricultural Research*, pp. 771-779, 2007. ↑42
- [56] FAO, "Pérdidas y desperdicio de alimentos en el mundo". Autor, 2011. ↑
- [57] J. Orjuela, I. Castañeda, J. Canal y J. Rivera, "La logística en la cadena de frutas". *Frutas y Hortalizas*, no 39, pp. 10 -15, 2015. ↑42
- [58] J. A. Orjuela Castro, L. E. Vidal Morales , A. A. Masmela Cortes y J. Rivera Velasco, "Logística en la cadena hortícola (Olerícola)". *Revista Frutas & Hortalizas*, vol. 41, no 1, Mayo-Junio 2015. ↑

---

### Lizeth Andrea Sanabria Coronado

Ingeniera Industrial, Universidad Distrital Francisco José de Caldas; integrante grupo de investigación en Cadenas de Abastecimiento Logística y Trazabilidad GICALyT.  
Correo electrónico: lizeth.asanabria@hotmail.com

---

### Andres Mauricio Peralta Lozano

Ingeniera Industrial, Universidad Distrital Francisco José de Caldas; integrante grupo de investigación en Cadenas de Abastecimiento Logística y Trazabilidad GICALyT.  
Correo electrónico: amperaltal@hotmail.com

---

### Javier Arturo Orjuela Castro

Ingeniero de Alimentos; ingeniero industrial; especialista en Ingeniería de Producción Universidad Distrital Francisco José de Caldas; magíster en Investigación de Operaciones y Estadística, Universidad Tecnológica de Pereira; PhD (C) en Ingeniería Industria y Organizaciones, Universidad Nacional de Colombia; docente e investigador en Logística, Universidad Distrital Francisco José de Caldas; líder grupo GICALyT.  
Correo electrónico: jorjuela@udistrital.edu.co

**Anexo I. Notación utilizada en el artículo**

Notación	Descripción
$x_j$	1 si la instalación es localizada en el nodo candidato $j \in J$ y 0 de otra manera
$x_h$	1 si la instalación es localizada en el nodo candidato $h \in J$ y 0 de otra manera
$y_{ij}$	1 si el nodo de demanda $i \in I$ es asignado al nodo de instalación candidato $j \in J$ y 0 de otra manera
$z_j$	1 si la estación de carga $j$ es elegida para operar
$Z_{jc}$	1 si el cargador tipo $c$ es asignado a la estación de carga $j$
$d_{ij}$	Distancia entre cada nodo de demanda $i \in I$ y cada instalación candidata $j \in J$
LMJ	La distancia mas corta desde el candidato estación de carga $j$ al molino de azucar
$p$	Número de instalaciones que serán abiertas
$w_i$	peso o demanda del punto $i$
$z_i$	1 si el nodo de demanda $i \in I$ es cubierto y 0 si no lo es
$t_{ij}$	Tiempo de transporte del punto $i$ hasta el punto $j$
$w_{ij}$	Cantidad de producto transferido desde el punto $i$ al punto $j$
$p_{ijp}$	Fracción de la demanda requerida por el punto $j$ del producto $p$ desde la locación $i$
$M$	numero maximo de estaciones de carga requeridas
$ms_{ij}$	Fracción de las cantidades enviadas desde el lugar de producción $i$ de ubicación centro $j$ (toneladas)
$md_{jh}$	Fracción de las cantidades enviadas desde la ubicación centro $j$ a la ubicación de consumo $h$ (toneladas)
$U_h, U_j$	Capacidad máxima de la instalación
$b_j$	Capacidad de servicio total de la estación de carga $j$ (toneladas/día)
$d_j$	Demanda de la instalación $j$
$\theta$	Coficiente de la velocidad de perecedero
$F$	Costo de la logística del sistema (transporte y costo de pérdida)
$c$	precio unitario de los productos perecederos
$Z$	Tiempo maximo de recorrido
$d$	Factor de descuento sobre los vinculos entre los centros
$t_{ik}+dT_{km}+T_{mj}$	Tiempo total de recorrido
$Y_{ikmj}, Y_{jm}, Y_{ij}$	Asignacion entre nodos $i, k, m, j$
$q_{ijp}$	Costo de proveer unidades $p$ desde el deposito $i$ hasta el nodo de demanda $j$
$g_{pj}$	Costo fijo por producto $p$ en la instalación $j$
$f_j$	Costo fijo de instalación $j$
$F_c$	Costo fijo de cargador tipo $c$ (\$)
$O_c$	Costo de operación de cargador tipo $c$ (\$/tonelada)
$h$	Costo de transporte (\$/km)
$b_c$	Capacidad de servicio de cargador tipo $c$ (ton/día)
$c_{ikmj}$	Costo de transportar entre el nodo origen $i$ , pasando por $k$ y $m$ , hasta el nodo terminal $j$
$\alpha$	Factor de proporción de envío entre nodos
$TPI_{ijn}$	Cantidad de movimientos de productos agrícolas frescos de la locación $i$ a la $j$ a través del modo de transporte $n$
$ppt$	Tasa de producto $p$ elegibles recogidas de la finca en el periodo $t$
$PP_i$	Cantidad de producto agrícola $p$ fresco recolectado dentro de la región $i$
$YPP_j$	0,1 si es abierta la plata en la locación $j$
$ppm_i$	Máxima cantidad de producto $p$ que pueden ser recolectadas dentro de la región

## Reporte de caso

## Una Mirada a la Web de los Datos. Caso de Estudio: Consumo de Servicios CKAN

*A View of the Web of Data. Case Study: Use of Services CKAN*

**Jhon Francined Herrera-Cubides<sup>1</sup>, Paulo Alonso Gaona-García\*<sup>1</sup>, Kevin Gordillo-Orjuela<sup>1</sup>**

<sup>1</sup>Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas. Bogotá-Colombia  
Correspondencia: pagaonag@udistrital.edu.co

Recibido: 05/07/2016. Modificado: 05/09/2016 Aceptado: 12/10/2016

### Resumen

**Contexto:** Se busca llevar a cabo el análisis, conexión y uso de los servicios ofrecidos por Comprehensive Knowledge Archive Network (CKAN, por sus siglas en inglés), con el fin de evaluar criterios base para obtener referentes preliminares de estudio sobre el estado de la web de los datos, a través de la exploración y acceso de los dataset publicados en el repositorio de datos abiertos DataHub.io.

**Método:** Empleamos los servicios ofrecidos por CKAN para la consulta y descarga de los dataset publicados en Datahub.io, para lo cual presentamos una serie de procesos llevados a cabo para analizar los datos descargados. La propuesta se compone de tres actividades clave: (1) revisión y análisis de las plataformas; (2) configuración y uso de los servicios prestados por la API; y (3) descarga y revisión de la información obtenida.

**Resultados:** Se configuraron y desplegaron los servicios requeridos, a través de CKAN, con el fin de llevar a cabo las consultas y descargas respectivas de dataset. Se procesó y analizó la información obtenida de los JSON descargados, permitiendo hacer un análisis preliminar comparativo, de la información obtenida acerca del comportamiento de la web de los datos.

**Conclusiones:** CKAN es una herramienta potente para gestionar catálogos de datos, permitiendo manejar una descripción de los datos y otra información relevante, tanto para las organizaciones que publican como para las personas que consultan dicha información, tales como categorías de organizaciones, formatos en que se encuentra disponible los datos, propietario de los datos, el tipo de licenciamiento de las publicaciones, enlaces a otros datos, entre otros datos, pertinentes para llevar a cabo un análisis de la web de los datos.

**Palabras clave:** CKAN, Datahub, JSON, Linked Data, web de datos.

**Idioma:** Español



Citación: J. F. Herrera, P. A. Gaona, K. Gordillo, "Una Mirada a la Web de los Datos. Caso de Estudio: Consumo de Servicios CKAN" INGENIERÍA, vol. 22, no. 1, pp. 46-64, 2017.

©Los autores; titular de derechos de reproducción Universidad Distrital Francisco José de Caldas. En línea DOI: <http://dx.doi.org/10.14483/udistrital.jour.reveng.2017.1.a03>

### Abstract

**Context:** In order to assess basic criteria so as to obtain preliminary guidelines on the current state of the Web of Data, we analyze the connection and use of the services offered by CKAN - Comprehensive Knowledge Archive Network; the analysis is conducted through exploration and connection to datasets published in the datahub.io open data repository.

**Method:** We use the services offered by CKAN for consultation and downloading datasets published in Datahub.io, we propose a procedure carried out to analyze the downloaded data. The proposal consists of three key activities: (1) review and analysis of platform, (2) Setting up and using the services provided by the API and (3) download and review of the information obtained.

**Results:** The required services offered by the platform CKAN were configured and deployed, in order to carry out queries and downloads related to each dataset. The obtained information was processed and analyzed from the downloaded JSON, allowing a comparative preliminary analysis of the information regarding the behavior of the Web of Data.

**Conclusions:** CKAN is a powerful tool to manage data catalogs. This tool can handle a description of the data and other relevant information, from organizations that publish to people who query such information. These queries provide information as categories of organizations, data formats and owners, the type of publication licenses, links to other data, among other which are relevant to perform an analysis of the Web data.

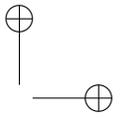
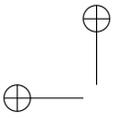
**Keywords:** CKAN, Web of Data, DataHub, JSON, Linked Data.

## 1. Introducción

La iniciativa propuesta por Tim Berners Lee [1] asociados a la vinculación de datos mediante Linked Data promete resolver los problemas asociados con el análisis y la interoperabilidad de los datos vinculados a recursos relacionados. En la actualidad existen varios ejemplos que permiten el uso de datos vinculados en diferentes áreas de conocimiento tales como la medicina, la geografía, la bibliotecología, entre otras. Por ejemplo, en el campo de la educación, la vinculación de los recursos educativos de diferentes repositorios de conocimiento es útil en la Web, permite el intercambio, la búsqueda y la navegación de objetos de aprendizaje [2]. Varios proyectos como Europeana [3], LinkedUp [4], [5], y Linked Education [2], han adoptado el enfoque de Linked Data y su objetivo responde a vincular conjuntos de datos educativos.

Por su parte, el proyecto Linked Open Data (LOD, por sus siglas en inglés) es otra iniciativa orientada a aplicar los principios de la web semántica con el fin de conectar recursos. LOD utiliza tecnologías como Resource Description Framework (RDF, por sus siglas en inglés) [6] y Uniform Resource Identifier (URI, por sus siglas en inglés), junto con un conjunto de principios denominados “datos vinculados”. Varias fuentes de datos, como Wikipedia, están ahora disponibles para los desarrolladores, que claramente se benefician de los conjuntos de datos vinculados con base en un modelo de datos común [3].

Datahub, sujeto de análisis en el presente artículo, es una plataforma para la gestión de datos, ba-



sada en CKAN - Comprehensive Knowledge Archive Network [17] - [18], que es una herramienta para la gestión y publicación de colecciones de datos (Datasets) en un ambiente web, utilizado por diferentes gobiernos, nacionales y locales, instituciones de investigación, entre otras (denominadas “organizaciones” en CKAN), que recogen una gran cantidad de datos. A través de sus servicios, los usuarios pueden buscar y encontrar los datos que necesitan.

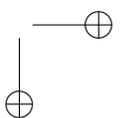
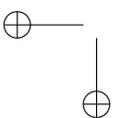
La motivación de este trabajo es presentar, a través de un estudio de caso basado en Datahub, cómo a partir de la conexión y el consumo de datos que pueda ofrecer el Application Programming Interface (API, por sus siglas en inglés), en este caso CKAN [17], se pueden obtener variables para la realización de un análisis preliminar sobre la visión de la web de los datos. Dicho caso estudio busca propiciar herramientas, para la identificación de tendencias en el modelo de la web de los datos, tomando como base datos históricos presentes en Datahub.io, como plataforma de gestión de datos, en el marco del proyecto de investigación sobre vinculación de datos que se viene adelantando.

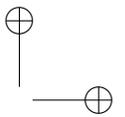
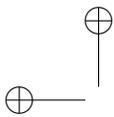
El resto del siguiente artículo se encuentra organizado de la siguiente manera: en la primera parte de se encuentra el estado del arte, donde se revisan los fundamentos de la temática explorada. Posteriormente, en la sección III se presenta el planteamiento metodológico utilizado para la exploración de los conjuntos de datos. En la sección IV, se presenta el desarrollo metodológico, junto con los hallazgos encontrados, en la sección V se presenta el análisis de resultados. Finalmente, en la sección VI se presentan las conclusiones y trabajo futuros.

## 2. Estado del arte

Gracias al modelo de web de los datos, varias fuentes de datos, como Wikipedia, están ahora disponibles para los desarrolladores, que claramente se benefician de los conjuntos de datos vinculados con base en un modelo de datos común [7]. Por su parte, el proyecto DBpedia [8] dispone de un conjunto de datos que actualmente se consideran como el eje central y más significativo entre los conjuntos de datos LOD [9]. En la web se encuentran iniciativas LOD tales como las plataformas de gestión de datos, entre ellas:

- Datahub (<https://datahub.io/>), objeto de examen en este artículo, la cual permite buscar y publicar datos, crear y gestionar grupos de Dataset, entre otras funcionalidades [10].
- Junar (<http://junar.com/>), plataforma de datos abiertos en la nube que facilita la publicación de datos por parte de gobiernos, empresas u otras organizaciones [11].
- Socrata (<https://www.socrata.com/>), plataforma escalable de publicación de datos en la nube que facilita la creación de iniciativas de datos abiertos sostenibles, ofreciendo un amplio conjunto de funcionalidades [11].
- data.gov.uk To Go (<http://guidance.data.gov.uk/>), kit del gobierno de Reino Unido para poner a disposición del público en general, su plataforma de publicación de datos de forma que cualquiera pueda desplegar una plataforma similar, preocupándose únicamente de adaptar la apariencia externa final [11].





- Plataforma de Gobierno Abierto OGoov (<http://www.ogoov.com/es/>), ofrece una serie de funcionalidades combinables entre sí según la orientación o iniciativas relacionadas con el gobierno abierto que se deseen priorizar: datos abiertos, transparencia y participación [11].

Estas entre otras plataformas, han permitido una mayor visibilidad de datos compartidos y han facilitado la participación de iniciativas como LOD Cloud [9], para la visualización de las organizaciones y proveedores de contenidos que han liberado y enlazado sus datos.

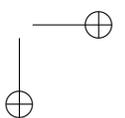
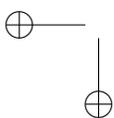
Como parte de un proyecto de investigación de tesis doctoral en curso, orientado en el dominio de la vinculación de datos, se plantea recabar información preliminar sobre el contenido de la nube LOD a 2016. Con tal fin, y dado que en experiencias como [9] donde se trabajó con DBPedia, para el presente estudio se propone analizar la plataforma de gestión de datos libres Datahub.io, de la Open Knowledge Foundation [12] – OKFN. Esta fundación tiene como visión que el conocimiento crea poder para muchos, no para unos pocos, los datos nos libera para tomar decisiones informadas sobre la forma en que vivimos, lo que compramos y quien recibe nuestro voto; la información y el conocimiento son accesibles, aparentemente, a todo el mundo. Por otra parte, Datahub.io, como una de las plataformas de gestión de datos y repositorio internacional de datos abiertos reconocidas [25], es una de las que más Dataset aporta a la conformación del Linked Open Data Cloud Diagram [9].

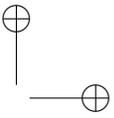
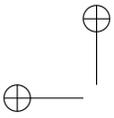
CKAN maneja un backend construido en Python, y un frontend construido en JavaScript. También usa el framework web Pylons y SQLAlchemy como ORM, con PostgreSQL como motor de base de datos. Tiene una arquitectura modular que permite desarrollar extensiones para proporcionar características adicionales, tales como harvesting o carga de datos, visualización de múltiples vistas, diferentes extensiones, un JSON API para leer, escribir y hacer consultas a los Dataset, soportado en más de 40 lenguajes, entre otras funcionalidades [23] - [24].

Al ser de código abierto, licenciado bajo términos GPL v3.0 Affero GNU, los usuarios CKAN pueden adaptar sus servicios, como lo han hecho, entre otros [13]:

- Africa's Largest Volunteer Driven Open Data Platform (<https://africaopendata.org/>).
- [data.amsterdam.nl](http://data.amsterdam.nl/) (<http://data.amsterdam.nl/>).
- Buenos Aires Data (<http://data.buenosaires.gob.ar/>).
- Paraguay Digital (<https://www.datos.gov.py/>).
- La plataforma cívica de datos abiertos de México (<http://datamx.io/>).
- Registro de conjuntos de datos abiertos en Noruega (<http://data.norge.no/>).

Ahora bien, para los propósitos de CKAN, los datos se publican en unidades denominadas "Dataset". Un Dataset es un paquete de datos —por ejemplo, podría ser las estadísticas de la delincuencia para una región, las cifras de gasto para un departamento gubernamental o las lecturas de temperatura de varias estaciones meteorológicas—. Cuando los usuarios buscan datos, los resultados de búsqueda que se obtiene son Dataset individuales. Un Dataset contiene dos componentes [13]:

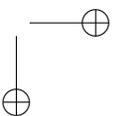
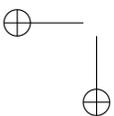




- Información o "metadatos" sobre los datos. Estos datos deben proveer la siguiente información:
  - a) Título, único a través de CKAN, de modo que sea breve pero específica. Por ejemplo: "densidad de población del Reino Unido según la región." es mejor que "Las cifras de población".
  - b) Descripción, información que la gente necesita saber cuándo se utilizan los datos.
  - c) Etiquetas, etiquetas que ayuden a la gente a encontrar los datos y la vinculan con otros datos relacionados.
  - d) Licencia, información de la licencia para que se sepa cómo se pueden utilizar los datos.
  - e) Organización, elegir quien es el propietario del Dataset.
- Un número de "recursos", que contienen los datos en sí. CKAN no le importa en qué formato están los datos (hoja de cálculo CSV (comma-separated values) o Excel, XML, PDF, archivo de imagen, RDF, entre otros). CKAN puede almacenar el recurso internamente o almacenar un enlace al recurso en sí, ubicado en otra parte en la web. Para los recursos se debe proveer la siguiente información:
  - a) Nombre, un nombre para el recurso.
  - b) Descripción, una breve descripción del recurso.
  - c) Formato, el formato de archivo del recurso, por ejemplo, CSV, XLS, JSON (JavaScript Object Notation), PDF, etc.
  - d) Visibilidad, un Dataset público puede ser visto por cualquier usuario del sitio. Un Dataset privado solo puede ser visto por los miembros de la organización propietaria del Dataset y no se mostrará en las búsquedas realizadas por otros usuarios.
  - e) Autor, el nombre de la persona u organización responsable de la producción de los datos.
  - f) Correo electrónico del autor.
  - g) Correo electrónico del responsable de mantenimiento.
  - h) Campos personalizados, si desea adicionar más datos.

Considerando lo anteriormente descrito, en Datahub.io, se registra 822 organizaciones (propietarios de los Dataset), creadas con previa autorización de un administrador, a través del envío de una solicitud. Estas organizaciones pueden crear, administrar y publicar colecciones de conjuntos de datos y pueden tener miembros tales como administradores (que añaden usuarios y gestionan la organización), y editores (que solo pueden añadir conjuntos de datos de la organización). Dicha distribución en organizaciones provee las siguientes características:

- Se proporciona una estructura de permisos y autorizaciones mucho más rica (en torno a la organización), que ofrece a los usuarios un mayor control sobre quién puede o no, editar y añadir conjuntos de datos.
- Proporciona una estructura orientada a la organización, para la presentación y la búsqueda de bases de datos.



- Ayuda a controlar problemas de spam, proporcionando más control sobre quién añade conjuntos de datos.

Dentro de las organizaciones incluidas en Datahub.io se encuentran las que se presentan en la tabla I, junto con el total de Dataset publicados por cada una de ellas.

**Tabla I.** Cantidad de dataset publicados por Organización [10]

Organizacion	Cantidad	%	Organizacion	Cantidad	%
Global	3418	36,8	Opendata.cz	46	0,5
London Datastore Archive	612	6,6	Civil Society	46	0,5
Países	1163	12,5	BudgIT Information Technology	43	0,5
Open Hampton Roads	531	5,7	Sport	38	0,4
Bio2RDF	378	4,1	IEE VIS	38	0,4
ie-ckan-net	272	2,9	Ayuntamiento de Zaragoza	37	0,4
It-ckan-net	263	2,8	Linked Education Cloud	36	0,4
Biportal	243	2,6	AKSW	33	0,4
no-ckan-net	232	2,5	Where Does My Money Go?	31	0,3
Linking Open Data Cloud	210	2,3	Bibliographic Data	30	0,3
leeds-datamill-archive	201	2,2	US State Spending and Revenue Data	29	0,3
LODCloud2014	196	2,1	Ontology Engineering	29	0,3
Economics Datasets	148	1,6	Urban Design Studio	26	0,3
cz-ckan-net	147	1,6	Tetherless World Constellation	26	0,3
OpenSpending	142	1,5	Negawatt Challenge	26	0,3
International Budget Partnership	83	0,9	eagle-i	25	0,3
DAL	79	0,9	School of Data	25	0,3
OWLG	63	0,7	OpenSpending Cameroon	25	0,3
ClimateData	54	0,6	Open Knowledge Brasil	24	0,3
Statutarni mesto Brno	48	0,5	DataID	23	0,2
Occupy	47	0,5	Wikimedia	22	0,2
Linking Open Data	47	0,5	International Food Policy Research Institute	22	0,2
			Energy Data	22	0,2

Como se observa en la tabla I, la categoría “Global” es la que maneja la mayoría de Dataset publicados. Lo anterior se debe al proceso de migración realizado por la plataforma, dejando en esta categoría a aquellos Dataset que no eran parte de ninguna organización. Por otro lado, se observa que la siguiente categoría de mayor publicación es la correspondiente a “países”. De igual forma, en Datahub.io se encuentran publicados 10.900 Dataset.

Es importante anotar que hay organizaciones que:

- No tienen ningún Dataset publicado.
- Tiene un archivo de prueba de carga, que no configurar ningún tipo de Dataset.
- Tiene uno o más de un Dataset publicado, en diferentes formatos (que van desde CSV hasta PDF; XML, Sparql - SPARQL Protocol and RDF Query Language, etc.), y con diferentes tipos de licenciamiento.
- Tiene publicados otros archivos diferentes a un Dataset.

Dentro de la literatura se encuentran algunos trabajos relacionados tales como [19], quienes describen el problema de los enlaces abiertos y la estrategia para resolver este problema. Por su parte, en [20] se muestra las cantidades de datos vinculados disponibles a partir de julio de 2009 y el número de enlaces entre conjuntos de datos RDF; y en [21] se presenta un estudio de estadísticas acerca de la estructura y contenido de la nube LOD.

Es así como a partir de este panorama, el siguiente documento tiene como propósito presentar un estudio de caso sobre la configuración, acceso y consumo de datos que pueda ofrecer el API CKAN [17], con el fin de poder consultar una muestra de los Dataset publicados en Datahub, seleccionados de forma aleatoria, con los cuales se podrá identificar y extrapolar, de manera preliminar, las tendencias del modelo de la web de los datos a partir de los datos históricos presentes en Datahub.io, como plataforma de gestión de datos. Análisis que se plantea en etapas posteriores del proyecto que enmarca esta investigación. El motivo de usar esta API se debe a que CKAN está escrito en Python y hace uso de una variedad de framework de código abierto, incluido Pylons, el cual es una combinación de varios framework de código abierto integrados que forman la base para aplicaciones de nivel empresarial basados en la web.

El almacenamiento y gestión de datos de CAKAN incluye el almacenamiento de archivos, gestión de metadatos, y la gestión de datos estructurados. Además, ofrece un mecanismo de plug-in, que permite a los desarrolladores extender rápidamente la funcionalidad de CKAN. De igual forma, CKAN posee la capacidad de soportar características geográficas, así como la exposición de metadatos de acuerdo con el catálogo estándar Open Geospatial Consortium (OGC, por sus siglas en inglés) y Catalog Service for the Web (CSW, por sus siglas en inglés). Por último, CKAN implementa funciones de limpieza cruciales tales como el logueo y gestión de usuarios [22].

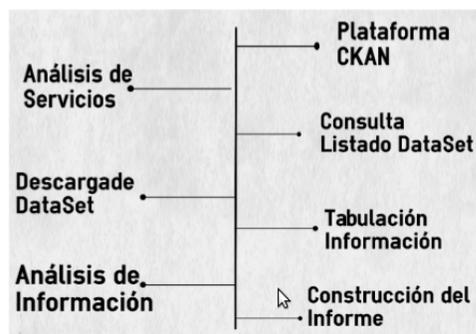


Figura 1. Metodología desarrollada para el proyecto

### 3. Planteamiento metodológico

Para llevar a cabo el proceso de conexión y descarga de los Dataset publicados en Datahub.io, se realizaron una serie de actividades tal como se describe en la figura 1.

Tal como se presenta en la figura 1 las actividades más relevantes son:

- Se revisó la plataforma CKAN, con el fin de establecer los servicios que ésta provee, al igual que la forma de acceder a ellos. Para esta labor se hizo uso del CKAN's API guide [14]- [15].
- Posteriormente se estudió la forma de conectarse a los servicios prestados por CKAN, con el fin de descargar los Dataset publicados en Datahub.io [14]- [15].

- Haciendo uso de los servicios de CKAN, se consultó el listado de los nombres de los Dataset publicados.
- Dada la cantidad de Dataset publicados, Datahub usa la estrategia de publicar grupos de Dataset (en sus respectivos formatos) por páginas, para un total de 545 páginas. Teniendo en cuenta la distribución anterior, se descargaron archivos JSON, formato de archivo que se descarga localmente, los cuales contienen los Dataset publicados por página en la plataforma de gestión de datos.
- Por último, con el fin de revisar la integridad de la descarga, se llevó a cabo un análisis comparativo de la información, haciendo uso de herramientas que permitieran visualizar la data descargada, con la revisión manual de algunos Dataset publicados.
- Construcción del presente artículo, de acuerdo a los resultados obtenidos.

## 4. Materiales y métodos

Dado que el presente documento se basa en la descripción de un caso de estudio acerca de la conexión y consumo de los servicios ofrecidos por CKAN y, con el fin de que el procedimiento seguido pueda ser replicado a *posteriori*, a continuación se describen el proceso desarrollado con el fin de poder consumir los servicios del API, y por ende consultar los Dataset, dispuestos en páginas, para posteriormente ser descargados en archivos JSON, los cuales permitirán llevar a cabo un análisis preliminar de las tendencias del modelo de la web de los datos. El proceso seguido, junto con los recursos necesarios para la conexión y consumo se describe a continuación.

### 4.1. Recursos para la conexión a Datahub.io – CKAN API:

Datahub.io, como sitio desarrollado usando la API de CKAN, permite hacer uso de las funcionalidades que provee la API como consultar los metadatos (información acerca del dato alojado),

**Tabla II.** Recursos utilizados en la Conexión a CKAN

Recurso	Descripción	Características del recurso
Ubuntu desktop o server 12.04 o superior	Distribución Linux, basada en Debian GNU/Linux. Sistema Operativo predominantemente enfocado a máquinas de escritorio, aunque también proporciona soporte para servidores	Las consultas realizadas en este proceso de exploración se realizaron usando Ubuntu 15.10
Python 2.7 o superior	Python es un lenguaje de programación de alto nivel, interpretado y de código abierto	Aunque se puede reemplazar con cualquier lenguaje que permita interactuar con el shell de Ubuntu. Para esta exploración se usó Python, dada su afinidad con Ubuntu
Visualizador de archivos .json	Archivos JSON – JavaScript Object Notation: formato estándar abierto que utiliza texto legible para transmitir objetos de datos que consisten en pares atributo-valor. Para visualizar este tipo de archivos, en Google Chrome o Firefox, se debe instalar el plugin JSONView. De igual forma se pueden visualizar con un editor que permita la edición de este tipo de archivos.	Para esta tarea se puede usar Sublime text 2/3 o cualquier editor que facilite la lectura
HTTPIE	Cliente HTTP de línea de Comando, utilizada para la comunicación e interacción entre usuarios y web Service. Provee un comando HTTP simple que permite realizar peticiones HTTP usando sentencias simples. Es usado para realizar pruebas, depuración e interacción con servidores HTTP. La mayoría de las distribuciones de Linux proporcionan un paquete que se puede instalar mediante el gestor de paquetes del sistema.	Proporciona un comando HTTP que permite el envío de peticiones HTTP arbitrarias utilizando una sintaxis simple y natural. Aunque HTTPIE es compatible con Python 2.6 y 2.7, se recomienda instalar HTTPIE en una versión superior de Python siempre que sea posible. Esto asegurará que algunas de las características HTTP más recientes, como SIN, funcionen de mejor forma.

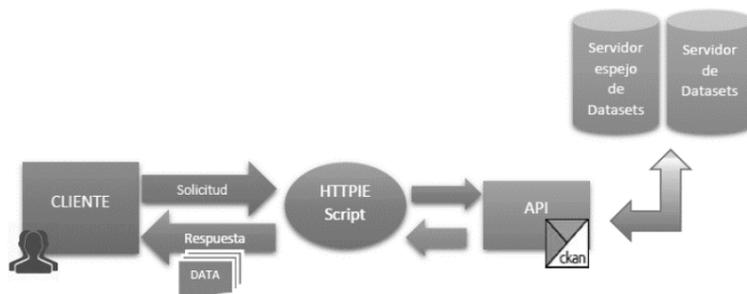


Figura 2. Interacción de Recursos de Conexión a CKAN

descargar la lista de los Dataset alojados, información de organizaciones, etiquetas que identifican determinado Dataset, entre otras funciones.

El proceso para realizar consultas sobre la API plantea el uso de los siguientes recursos (Tabla II).

La figura 2 se muestra la interacción de los recursos necesarios para conectarse a los servicios de CKAN.

#### 4.2. Proceso de instalación y consulta al API:

Instalación HTTPie: una vez ubicado en el entorno Ubuntu, se procede a abrir una terminal y ejecutar el comando:

```
sudo apt-get install httpie
```

La figura 3 muestra los resultados de la ejecución del citado comando

```
kevin@kevin-Lenovo-U310: ~  
kevin@kevin-Lenovo-U310:~$ sudo apt-get install httpie  
Leyendo lista de paquetes... Hecho  
Creando árbol de dependencias  
Leyendo la información de estado... Hecho  
httpie ya está en su versión más reciente.  
0 actualizados, 0 nuevos se instalarán, 0 para eliminar y 160 no actualizados.  
kevin@kevin-Lenovo-U310:~$
```

Figura 3. Instalación de HTTPie

Una vez ha terminado el proceso de instalación, ya se pueden realizar peticiones a servidores http, en este caso CKAN. La documentación de HTTPie se puede encontrar en el repositorio <https://github.com/jkbrzt/httpie>.

Ahora bien, la estructura básica para una consulta HTTPie se formula de la siguiente manera:

```
http [Flags] [Method] URL [Item [Item]]
```

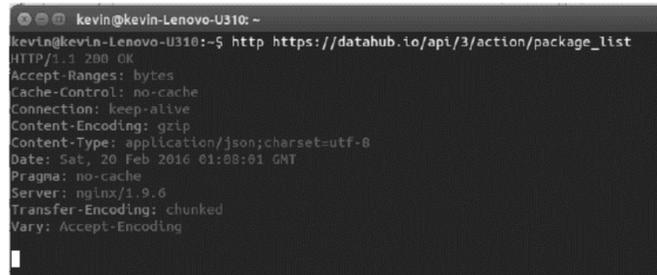
Consulta de los DataSet: CKAN provee métodos para consultar los repositorios del sitio que ha sido desarrollado con su API. Para hacer uso de dichos métodos basta con agregar al sitio principal (que haga uso de CKAN) la sintaxis:

`/api/3/action/ [método del CKAN]`

Con el fin de aclarar lo descrito anteriormente, se hará una consulta a Datahub.io de todos los nombres de los Dataset que tiene y, por defecto, el servidor devolverá el resultado en una lista con estructura tipo JSON. Para realizar este proceso se abre una terminal y se digita:

`http https://datahub.io/api/3/action/package_list`

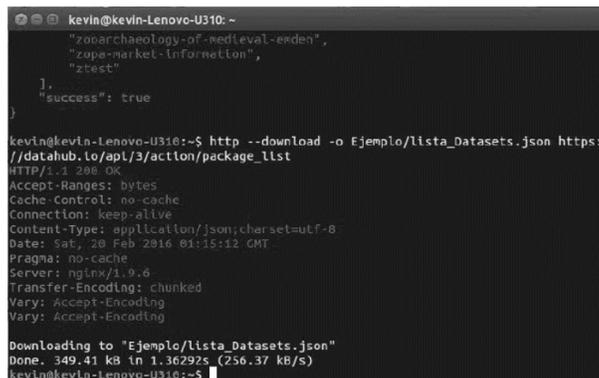
La figura 4 muestra los resultados de la ejecución del citado comando.



```
kevin@kevin-Lenovo-U310: ~$ http https://datahub.io/api/3/action/package_list
HTTP/1.1 200 OK
Accept-Ranges: bytes
Cache-Control: no-cache
Connection: keep-alive
Content-Encoding: gzip
Content-Type: application/json;charset=utf-8
Date: Sat, 20 Feb 2016 01:08:01 GMT
Pragma: no-cache
Server: nginx/1.9.6
Transfer-Encoding: chunked
Vary: Accept-Encoding
```

Figura 4. Consulta del Listado de Nombres de DataSet

Seguido a esta ejecución, se desplegará dentro de la misma consola los nombres de todos los Dataset de Datahub.io. Si adicionalmente se quiere descargar la información consultada y volcarla a un archivo, se debe agregar las banderas `--download` y `-o`, incluidas en HTTPie. La bandera `--download` permite descargar el archivo; y la bandera `-o` permite nombrar y ubicar el archivo descargado. Por ejemplo, para descargar la lista anterior se realiza el proceso descrito en la figura 5.



```
kevin@kevin-Lenovo-U310: ~$ http --download -o Ejemplo/lista_Datasets.json https://datahub.io/api/3/action/package_list
HTTP/1.1 200 OK
Accept-Ranges: bytes
Cache-Control: no-cache
Connection: keep-alive
Content-Type: application/json;charset=utf-8
Date: Sat, 20 Feb 2016 01:15:12 GMT
Pragma: no-cache
Server: nginx/1.9.6
Transfer-Encoding: chunked
Vary: Accept-Encoding
Vary: Accept-Encoding

{"dataset": ["zooarchaeology-of-medieval-eden", "zoo-market-information", "ztest"], "success": true}
kevin@kevin-Lenovo-U310: ~$ http --download -o Ejemplo/lista_Datasets.json https://datahub.io/api/3/action/package_list
HTTP/1.1 200 OK
Accept-Ranges: bytes
Cache-Control: no-cache
Connection: keep-alive
Content-Type: application/json;charset=utf-8
Date: Sat, 20 Feb 2016 01:15:12 GMT
Pragma: no-cache
Server: nginx/1.9.6
Transfer-Encoding: chunked
Vary: Accept-Encoding
Vary: Accept-Encoding

Downloading to "Ejemplo/lista_Datasets.json"
Done, 349.41 kB in 1.36292s (256.37 kB/s)
```

Figura 5. Descarga de la Consulta de Nombres de DataSet

Como resultado de esta ejecución se descarga el archivo con el nombre “lista\_Datasets.json”, en el folder “Ejemplo”, previamente creado.

Por otro lado, para descargar todos los Dataset de Datahub.io, se puede hacer uso del método *current\_package\_list\_with\_resources*. Sin embargo, este método solo descarga de a diez Dataset si no se le especifica algún tipo de parámetro. Adicionalmente, su capacidad máxima de descarga de Dataset es de 1000, es decir que el archivo de descarga tendrá como máximo 1000 Dataset. Por ende, para descargar los Dataset publicados en la plataforma se escribe un script en Python, que haga uso de un ciclo *for*, que permita descargar los Dataset completos de cada página hasta completar 535 páginas. En la figura 6 se muestra el script y el resultado de su ejecución.

```
script.py *
import http
import os

for i in range(1,535):
    os.system("http --download -o MetaData_DataHub/Datasets_Pagina"+str(i)+".json https://datahub.io/api/3/action/
current_package_list_with_resources limit=20 page="+str(i)+"")
```

Figura 6. Archivo Python para descarga de DataSet

Para ejecutar el script desde la terminal, se debe ubicar en el directorio donde se va a guardar el archivo y se ejecuta el comando Python [Nombre del archivo], que para este caso corresponde: Python script.py

El resultado de la iteración es la generación de un archivo por página (535 archivos), como se muestra en la figura 7.

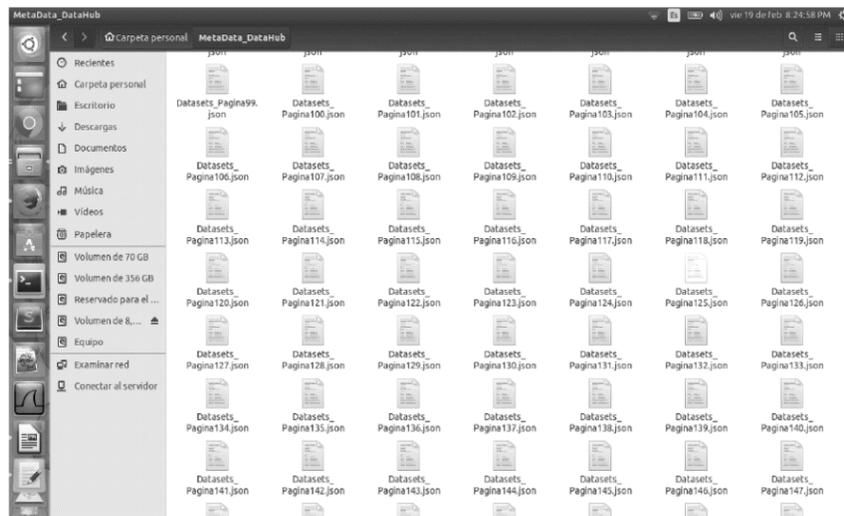


Figura 7. Archivos JSON descargados

## 5. Resultados y discusión

Como primer resultado obtenido en el proceso de conexión, se descargó el listado de Dataset publicados en la plataforma. La figura 8 muestra, al costado derecho el archivo JSON resultante de la consulta y al costado izquierdo el árbol de consulta resultante, donde cada valor de arreglo “result” corresponde a un nombre de Dataset consultado.

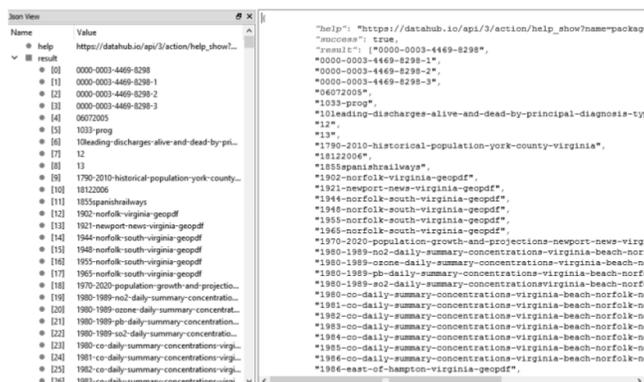


Figura 8. Listado de DataSet consultados

En total se obtuvieron 10694 entradas a la variable “result”, que representa la misma cantidad de nombres de Dataset consultados.

Posterior a esta consulta, se procedió a la descarga del listado de Dataset, como se especificó en la sección de Métodos y Materiales, generando 545 archivos correspondientes al mismo número de páginas existentes en la plataforma. Con este producto obtenido, se procedió a comparar y verificar la información resultante del proceso de descarga:

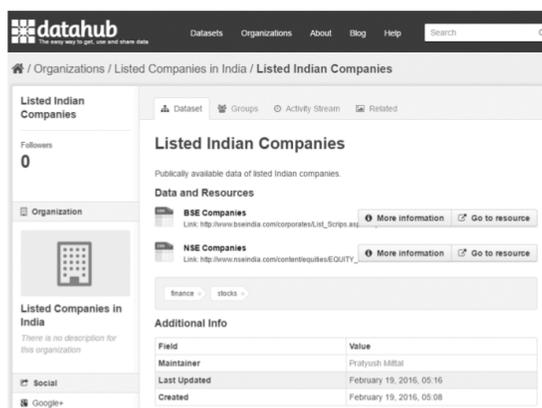


Figura 9. Organización consultada en Datahub.io

- a. DataSet consultado en la plataforma de Datahub.io: para revisar manualmente el proceso se tomó como ejemplo y punto de referencia la Organización “Listed Indian Companies”, la cual tiene un DataSet publicado en dos formatos, como se observa en la figura 9.
- b. En la figura 10 se observa el segmento del archivo JSON, que contiene la información de esta organización.

```

1  {
  "help": "https://datahub.io/api/3/action/help_show?name=current_package_list_with_resources",
  "success": true,
  "result": [
    {
      "license_title": "Creative Commons Attribution",
      "maintainer": "Pratyush Mittal",
      "relationships_as_object": [],
      "private": false,
      "maintainer_email": "pratyushmittal@gmail.com",
      "num_tags": 2,
      "id": "b78af181-811b-4182-b624-82919ea05556",
      "metadata_created": "2016-02-19T05:08:41.492537",
      "metadata_modified": "2016-02-19T05:16:00.401035",
      "author": "",
      "author_email": "",
      "state": "active",
      "version": "",
      "creator_user_id": "ede20aa6-dda4-4535-a82e-58ff4d48e9930",
      "type": "dataset",
      "resources": [
        {
          "mimetype": null,
          "cache_url": null,
          "hash": "",
          "description": "Link: http://www.bseindia.com/corporates/List_Scripts.aspx?expandable=1",
          "name": "BSE Companies",
          "format": "CSV",
          "url": "https://datahub.io/dataset/b78af181-811b-4182-b624-82919ea05556/resource/372c614e-064f-4521-8136-068366be578c/download/istofscrips.csv",
          "cache_last_updated": null,
          "package_id": "b78af181-811b-4182-b624-82919ea05556",
          "created": "2016-02-19T05:10:47.980032",
          "state": "active",
          "mimetype_inner": null,
          "webstore_last_updated": null,
          "last_modified": null,
          "position": 0,
          "revision_id": "1eb09546-0ee3-4089-a2b8-96b5b55aaace",
          "webstore_url": null,
          "url_type": "upload",
          "id": "372c614e-064f-4521-8136-068366be578c",
          "resource_type": null,
          "size": null,
          "mimetype": null,
          "cache_url": null,
          "hash": "",
          "description": "Link: http://www.nseindia.com/content/equities/EQUITY_L.csv",
          "name": "NSE Companies",
          "format": "CSV",
          "url": "https://datahub.io/dataset/b78af181-811b-4182-b624-82919ea05556/resource/ab8acfa-323f-433c-9388-3ba1328fc490/download/equity1.csv",
          "cache_last_updated": null,
          "package_id": "b78af181-811b-4182-b624-82919ea05556",
          "created": "2016-02-19T05:11:40.260386",
          "state": "active",
          "mimetype_inner": null,
          "webstore_last_updated": null,
          "last_modified": null,
          "position": 1,
          "revision_id": "30f94f18-0df5-4d75-9470-b91f6007614",
          "webstore_url": null,
          "url_type": "upload",
          "id": "ab8acfa-323f-433c-9388-3ba1328fc490",
          "resource_type": null,
          "size": null,
          "num_resources": 2,
          "tags": [
            {
              "vocabulary_id": null,
              "state": "active",
              "display_name": "finance",
              "id": "98b2affc-d7f0-4cfd-b980-1339b805873b",
              "name": "finance",
              "vocabulary_id": null,
              "state": "active",
              "display_name": "stocks",
              "id": "152a3b6f-f347-41af-8de6-535e51c28af6",
              "name": "stocks",
              "groups": []
            }
          ],
          "license_id": "cc-by",
          "relationships_as_subject": [],
          "organization": {
            "description": "",
            "created": "2016-02-12T07:56:03.046042",
            "title": "Listed Companies in India",
            "name": "listed-indian-companies",
            "is_organization": true,
            "state": "active",
            "image_url": "",
            "revision_id": "ff247339-2560-4c75-a5b4-e2191af950a8",
            "type": "organization",
            "id": "221a69a5-5609-431e-b246-acc2b4facf70",
            "approval_status": "approved",
            "name": "listed-indian-companies",
            "isopen": true,
            "url": "",
            "notes": "Publicly available data of listed Indian companies.",
            "owner_org": "221a69a5-5609-431e-b246-acc2b4facf70",
            "extras": [],
            "license_url": "http://www.opendefinition.org/licenses/cc-by",
            "title": "Listed Indian Companies",
            "revision_id": "9e96eee-a26e-4059-8b6e-cc130e2308fb",
            "license_title": "Creative Commons Attribution",
            "maintainer": "Heather Mann",
            "relationships_as_object": [],
            "private": false,
            "maintainer_email": "heather.mann@gmail.com",
            "num_tags": 4,
            "id": "b8321a70-4fa0-422f-841f-f76e99a4743a",
            "metadata_created": "2016-02-18T19:01:41.047555",
            "metadata_modified": "2016-02-18T19:30:19.946115",
            "author": "Heather Mann, Ximena Garcia-Rada, Lars Hornuf, Juan Tafur, Dan Ariely",
            "author_email": "heather.mann@gmail.com",
            "state": "active",
            "version": "",
            "creator_user_id": "e59b6d7b-294e-4f74-b260-09952249160",
            "type": "dataset",
            "resources": [
              {
                "mimetype": null,
                "cache_url": null,
                "hash": "",
                "description": "",
                "name": "Cross-cultural dishonesty study - Reports of roll outcomes on die task",
                "format": "CSV",
                "url": "https://datahub.io/dataset/b8321a70-4fa0-422f-841f-f76e99a4743a/resource/4f754a6-2fd-4824-9f6e-5ec23a682018/download/cross-cultural-dishones"
              }
            ]
          }
        }
      ]
    }
  ]
}

```

Figura 10. Segmento del archivo JSON descargado en la consulta

- c. Haciendo uso de herramientas visualizadoras de archivos JSON, se puede revisar la información contenida en el segmento JSON, tal como los muestran las figuras 11 y 12.
- d. De igual forma, con este tipo de herramientas se pueden revisar parámetros de evaluación, tales como tipo de licenciamiento, última fecha de actualización, tipo de formatos publicados, entre otros que pueden brindar información relevante al momento de evaluar el estado de la web de los datos. La figura 13 permiten evidenciar la información brindada por este tipo de herramientas.

help	https://datahub.io/api/3/action/help_show?name=current_package_list_with_resources					
success	true					
result (20)						
	license_title	maintainer	relationships_a...	private	maintainer_email	num_tags
1	Creative Commons Attribution	Pratyush Mittal	relationships_a...	false	pratyushmittal@gmail.com	2
2	Creative Commons Attribution	Heather Mann	relationships_a...	false	heather.mann@gmail.com	4

Figura 11. Visualización del archivo JSON en modo cuadrícula



Figura 12. Visualización del archivo JSON en modo árbol

	( ) mimetype	( ) cache_url	( ) hash	( ) description	( ) name	( ) format	( ) url
1	null	null		Link: <a href="http://www.bseindia.com/corporates/List_Scripts.aspx?expandable=1">http://www.bseindia.com/corporates/List_Scripts.aspx?expandable=1</a>	BSE Companies	CSV	<a href="https://datahub.io/datasets/b78af181-811b-4182-b624-82919ea05556/resource/372c614e-064f-4521-8136-068368be578c/download/istofsc_rips.csv">https://datahub.io/datasets/b78af181-811b-4182-b624-82919ea05556/resource/372c614e-064f-4521-8136-068368be578c/download/istofsc_rips.csv</a>
2	null	null		Link: <a href="http://www.nseindia.com/content/equities/EQUITY_L.csv">http://www.nseindia.com/content/equities/EQUITY_L.csv</a>	NSE Companies	CSV	<a href="https://datahub.io/datasets/b78af181-811b-4182-b624-82919ea05556/resource/ab8acf8a-323f-433c-9388-3ba1328fc490/download/equity-1.csv">https://datahub.io/datasets/b78af181-811b-4182-b624-82919ea05556/resource/ab8acf8a-323f-433c-9388-3ba1328fc490/download/equity-1.csv</a>

Figura 13. Información obtenida de los DataSet

Como se observa en la figura 13, el Dataset consultado como ejemplo de verificación, tiene dos recursos publicados: BSE Companies y NSE Companies, ambos en formato CSV, bajo licencia cc-by, con fecha de última actualización el 19 de febrero de 2016.

- e. Para la realización de un análisis preliminar de la información descargada, se seleccionó, de forma aleatoria, una muestra de 54 organizaciones y se procedió a revisar los Dataset publicados, 311 en total, con el fin de levantar la información acerca de dominios, formatos licenciamiento y última fecha de publicación de los Dataset.

Como resultado de este proceso se obtuvo la siguiente información:

• **Dominio de Datos:** En los 311 Dataset revisados, en la figura 14 se categorizan los dominios de la siguiente manera:

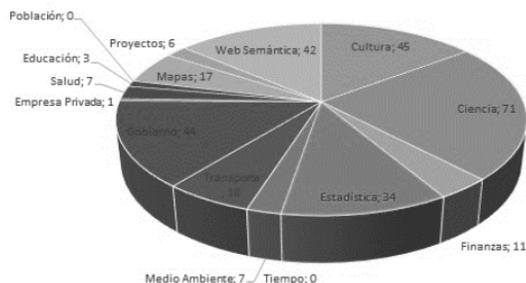


Figura 14. Dominio de Datos versus Cantidad de DataSet

En la revisión de dominio de publicación se encuentra que los dominios de tipos de datos abiertos con mayor publicación son:

- Ciencia, con un 23 %, en la cual los mayores subdominios son arqueología, química y bioinformática.
- Cultura, gobierno y web semántica, con un 14 %.
- Estadística, con un 11 %.

• **Formatos de publicación:** como se observa en la revisión preliminar realizada en Datahub.io (Tabla III), los DataSet están cargados en la plataforma en diferentes formatos (algunos de las organizaciones cargan más de un tipo de formato).

En la revisión de estos formatos, se puede observar que la plataforma no maneja un estándar unificado en cuanto a los formatos subidos, lo anterior añadido a que los formatos que se cargan no presentan una estandarización, lo que se puede observar al encontrarse como XLS, XLSX, Hoja de Cálculo, Excel, SpreadShet, MS Excel.

De igual forma, se identifica que hay publicaciones de Dataset realizadas en formatos no estructurados como PDF, PNG. Como se observa en la tabla III, en Datahub.io se pueden cargar diferentes formatos de un Dataset, encontrando por ejemplo que hay 1389 publicaciones que tiene formato de hojas de cálculo en sus formatos de publicación, lo que representa que un 13 % de los Dataset publicados y, como formatos propietarios, requieren herramientas que no son públicas.

También se identifica que varias organizaciones le han apostado a depurar sus publicaciones, escalando en el modelo propuesto por Berners-Lee [1], ofreciendo datos en formatos estructurados, abiertos al público y que permiten llevar a cabo el proceso de vinculación. Sin embargo, en este proceso también se observa en Datahub, situaciones como la disparidad de formatos de publicación, algunos de los cuales no permiten la interoperabilidad semántica, limitándose a la descripción sintáctica de la información o, por el contrario, publicando formatos que dificultan el procesamiento por las máquinas.

**Tabla III.** Cantidad de formatos de publicación, obtenidos de la muestra analizada

Formato	Cantidad	%	Formato	Cantidad	%
CSV	1049	11,4	application/zip	101	1,1
XLS	803	8,7	example/turtle	95	1,0
Example rdf+xml	750	8,1	XML	92	1,0
Api/sparql	674	7,3	Json	90	1,0
text/csv	403	4,4	GeoJSON	90	1,0
meta/void	356	3,9	TopoJSON	88	1,0
application/rdf+xml	337	3,6	application/octet-stream	87	0,9
Csv	319	3,5	application/json	81	0,9
PDF	298	3,2	Xml	72	0,8
XLSX	251	2,7	RDF	68	0,7
ODS	248	2,7	application/x-ntriples	66	0,7
HTML	224	2,4	Pdf	58	0,6
Html	215	2,3	rdf/turtle	57	0,6
Aspx	210	2,3	Other	53	0,6
application/pdf	190	2,1	Data File in EXCEL	49	0,5
Spreadsheet	185	2	ZIP	48	0,5
text/turtle	176	1,9	JPG	48	0,5
meta/rdf-schema	165	1,8	Shp	46	0,5
text/html	157	1,7	JSON	46	0,5
SHP	144	1,6	Zip	45	0,5
application/vnd.ms-ex-cel	140	1,5	application/trig	44	0,5
KML	132	1,4	PNG	42	0,5
meta/sitemap	123	1,3	mapping/owl	411	0,4
Xls	101	1,1	example/rdfa	40	0,4
			GML	39	0,4

• **Fecha de última actualización:** en cuanto a la fecha de última actualización, teniendo en cuenta que bajo los Principios de Open Data [16], estos datos deben ser “Actualizados”, con el fin de que no pierdan su valor y sean precisos, como resultado de la exploración muestral realizada, se identifica que el 1 % de los mismos alcanzaría un grado de actualización frecuente, un 7,7 % un grado de actualización media, mientras que un 60 % presenta un grado de actualización deficiente.

En este aspecto es importante resaltar que varios de los dominios corresponden a datos estadísticos generados de periodos anteriores, los cuales no requerirían de una actualización permanente, sin embargo, esta disparidad en el proceso de actualización permite reflexionar acerca del grado de actualización de los Dataset publicados.

## 6. Conclusiones y trabajos futuros

CKAN es una herramienta potente para gestionar catálogos de datos, permitiendo manejar una descripción de los datos y otras informaciones relevantes, tanto para las organizaciones que publican como para las personas que consultan dicha información, tales como categorías de organizaciones, formatos en que se encuentra disponible los datos, propietario de los datos, el tipo de licenciamiento de las publicaciones, enlaces a otros datos.

Por otro lado, CKAN es una herramienta usada en muchos catálogos de datos abiertos, disponibles en Web. Las organizaciones que hacen uso de CKAN, tanto privadas como públicas, publican

sus datos haciendo uso de algún nivel de las recomendaciones de buenas prácticas de Open Data y las propuestas por Berners-Lee.

En cuanto a los servicios ofrecidos por CKAN, para la consulta y descarga de la data, son servicios consistentes, que ofrecen peticiones y respuestas en formato JSON, sin restricciones al público. Existen herramientas tales como Python, Java, Ruby, PHP, que pueden interactuar con el API REST de CKAN, permitiendo obtener la información solicitada, la cual puede ser visualizada en diferentes formatos, haciendo uso de herramientas visualizadoras, que permiten hacer las consultas respectivas de formas más ágiles.

Como lo plantea [24], es complejo seleccionar una plataforma sin realizar un estudio cuidadoso de las necesidades de los interesados. En el estudio citado, se plantea que CKAN se está convirtiendo en uno de los referentes en cuanto a la gestión de catálogos de fuentes de datos, pero que a su vez presenta desventajas como el mantenimiento de una plataforma tecnológica diferente al CMS usado para gestionar los datos, falta de soporte nativo de RDF para el enriquecimiento semántico de los datos, y el despliegue de una aplicación web diferente al portal CMS usado [26].

Por su parte, Datahub.io, como plataforma de gestión de datos, permite una publicación de datos ágil, y con pocas restricciones al usuario final, a través de una interfaz web sencilla. Por otro lado, si como desarrollador o investigador, se requiere a Datahub.io, para consultar o descargar sus Dataset, ofrece una como apoyo el CKAN API, interface que interactúa con diferentes herramientas, ofrece una variedad de servicios, para llevar a los procesos requeridos.

Ahora bien, dentro de las situaciones que se observan en las publicaciones realizadas en Datahub.io, corresponde a la interpretación que algunas organizaciones en Datahub le han dado al concepto de Datos Abiertos, tomando a la plataforma como una estrategia para publicar o promocionar escritos, artículos, eventos, etc., que no poseen ningún tipo de vinculación ni de contexto ampliado a los datos expuestos. Tal es el caso de la organización 'Mente Clara', donde sus 13 publicaciones son imágenes y PDF (artículos) acerca del budismo tántrico tibetano.

Otro aspecto identificado es el acceso a ciertos recursos de los Dataset tomados como muestra, dado que las URI no accedían, generando inconvenientes para el proceso de vinculación. Este tipo de problemas podría ser causado a la falta de actualización de los datos publicados en la plataforma, entre otras situaciones. Independiente del origen de dicho problema, uno de los factores claves de la web de los datos es el proceso de identificación y vinculación de datos que aporten contexto al dato publicado, y la falta de enlaces resolubles afecta el desarrollo de esta tecnología.

De otro lado, en el caso de los gobiernos, por ejemplo, se identifican publicaciones que responden a los principios de transparencia y apertura de información, pero que bajo los principios de Datos Abiertos y de Linked Data, no aportan contexto a la información expuesta, limitándose a publicar listados de cuadros en hojas de cálculo, algunos de los cuales no se actualizan de forma periódica. Un ejemplo de ello es por ejemplo la organización 'Bolivia' que registra seis Datasets, en los cuales todos tiene como última fecha de actualización hace en promedio dos años, y la información contenida en ellos corresponden a tablas de hojas de cálculo de informes de indicadores sociales, que no se complementan o añaden contexto a la información, con otros enlaces o datos de interés.

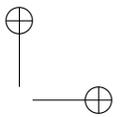
Como trabajos futuros del proceso de investigación, y dados los hallazgos identificados, se plantea: 1) llevar a cabo un análisis más detallado del estado de la Web de los datos, a través de herramientas de analítica visual, 2) realizar un análisis del estado de la Web de los datos en el área de recursos digitales abiertos mediante propuestas de accesos seguros a plataformas LCMS [27], objeto de estudio de la investigación, y por último, 3) el planteamiento de una propuesta metodológica para procesos de vinculación de datos de cara a los recursos digitales abiertos, analizando los hallazgos encontrados para su aplicación en entornos educativos.

### Agradecimientos

Esta investigación se lleva a cabo en el marco de la formación doctoral en Ingeniería, en la Universidad Distrital Francisco José de Caldas. De igual forma, la temática planteada se configura como una línea de investigación de Grupo GIIRA.

### Referencias

- [1] T. Berners-Lee, C. Bizer, T. Heath, “Linked data-the story so far”. *International Journal on Semantic Web and Information Systems*, vol. 5, pp. 1-22, 2009. ↑ 47, 60
- [2] S. Dietze, H. Yu, D. Giordano, E. Kaldouidi, N. Dovrolis, D. Taivi, “Linked Education: Interlinking Educational Resources and the Web of Data”. *27th annual ACM symposium on Applied Computing*, pp. 366-371, 2012. [En línea]. Disponible en: <http://oro.open.ac.uk/31077/>. ↑ 47
- [3] B. Haslhofer, A. Isaac, “Data. europea. eu: The European Linked Open Data Pilot ”. *International Conference on Dublin Core and Metadata Applications*, pp. 94-104, 2011. [En línea]. Disponible en: <http://depapers.dublincore.org/pubs/article/view/3625/1851>. ↑ 47
- [4] D’Aquin, M. et al ., “Building the Open Elements of an Open Data Competition ” *D-Lib Magazine*, vol. 20, p. 3, 2014. [En línea]. Disponible en: <http://www.dlib.org/dlib/may14/daquin/05daquin.html>. ↑ 47
- [5] L. Project, “Linking Web Data for Education”. [En línea]. Disponible en: <http://linkedup-project.eu/>. ↑ 47
- [6] G. Klyne, J. Carroll, “Resource Description Framework (RDF): Concepts and Abstract Syntax ”. 2006. Edited by Brian McBride. [En línea]. Disponible en: <http://www.citeulike.org/group/2170/article/532408>. ↑ 47
- [7] M. Hausenblas, “Exploiting Linked Data To Build Web Applications ”. *IEEE Internet Computing*, vol. 13, no 4, p. 68. 2009. ↑ 48
- [8] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, Z. Ives, “DBpedia: A Nucleus for a Web of Open Data ”. *The Semantic Web , Busan, Korea, Springer, 2007*. ↑ 48
- [9] LOD-Cloud. “The Linking Open Data Cloud Diagram”. S.f. [En línea]. Disponible en: <http://lod-cloud.net/>. ↑ 48, 49
- [10] Datahub, “Datahub Project”. S. f. [En línea]. Disponible en: <http://datahub.io/dataset?tags=lod>. ↑ 48, 51
- [11] Ministerio de Hacienda y Administración Pública, Ministerio de Industria Energía y Turismo, “Plataformas de publicación de datos abiertos”. [En línea]. Disponible en: <http://datos.gob.es/sites/default/files/informe-herramientas-publicacion.pdf>. ↑ 48, 49
- [12] OKF, “Open Knowledge Foundation ”. S. f. [En línea]. Disponible en: <http://services.okfn.org/>. ↑ 49
- [13] CKAN, “The Open Source Data Portal Software”. S. f. [En línea]. Disponible en: <http://ckan.org/>. ↑ 49
- [14] CKAN, “API Guide-CKAN’s API for developers”. S. f. <http://docs.ckan.org/en/latest/api/index.html>. ↑ 52
- [15] CKAN, “Ckan Wiki-CKAN Pages”. S.f. [En línea]. Disponible en: [https://github.com/ckan/ckan/wiki/\\_pages](https://github.com/ckan/ckan/wiki/_pages). ↑ 52
- [16] J. Wonderlich, “Ten Principles for Opening up Government Information”. 2010. [En línea]. Disponible en: <http://sunlightfoundation.com/policy/documents/ten-open-data-principles/>. ↑ 61
- [17] CKAN, “CKAN API Guide ”. S. f. [En línea]. Disponible en: <http://docs.ckan.org/en/latest/api/>. ↑ 48, 52
- [18] J. Winn, “Open Data and the Academy: an Evaluation of CKAN for Research Data Management. (IASSIST 2013)”. 28-31 Mayo 2013. [En línea]. Disponible en: <http://eprints.lincoln.ac.uk/9778/1/CKANEvaluation.pdf>. ↑



48

- [19] E. Rajabi, S. Sánchez-Alonso, M.-A. Sicilia, "Analyzing broken links on the web of data: An experiment with DBpedia". *Journal of the Association for Information Science and Technology*, vol. 65, no 8, p. 1721–1727, 2014. [Online]. Disponible en: <http://onlinelibrary.wiley.com/doi/10.1002/asi.23109/abstract>. ↑ 52
- [20] C. Bizer, "The Emerging Web of Linked Data". *IEEE Intelligent Systems*, vol. 24, no 5, pp. 87-92, 2009. [Online]. Disponible en: <http://lpis.csd.auth.gr/mtpx/sw/material/IEEE-IS/IS-24-5.pdf>. ↑ 52
- [21] HPI Institut, "State of LOD Cloud". 2011. [En línea]. Disponible en: <http://lod-cloud.net/state/>. ↑ 52
- [22] M. Allison, S. Richard, K. Patten, C. Caudill-Daugherty, A. Anderson, "Open Access to Geoscience Data for Exploration and Assessment". 19 al 25 Abril 2015. [En línea]. Disponible en: <http://www.geothermal-energy.org/pdf/IGASTandard/WGC/2015/33032.pdf>. ↑ 52
- [23] W. Mao, J. Jan, "Visualization of Open Data: a Case Study of Climate Data". 36 Asian Conference on Remote Sensing. 19-23 de Octubre de 2015. Manila, Philippines. [En línea]. Disponible en: <http://www.acrs2015.org/list-of-accepted-abstracts/>. ↑ 49
- [24] R. Carvalho, J. Aguiar, J. Rocha, C. Ribeiro, "A Comparison of Research Data Management Platforms: Architecture, Flexible Metadata and Interoperability". Junio de 2016. [En línea]. Disponible en: [https://www.researchgate.net/publication/303918099\\_A\\_comparison\\_of\\_research\\_data\\_management\\_platforms\\_architecture\\_flexible\\_metadata\\_and\\_interoperability](https://www.researchgate.net/publication/303918099_A_comparison_of_research_data_management_platforms_architecture_flexible_metadata_and_interoperability) ↑ 49, 62
- [25] CURE, "Infraestructura semántica basada en el paradigma de datos abiertos para la gestión de investigación de las universidades españolas". CRUE Universidades Españolas, 2016. [En línea]. Disponible en: <http://tic.crue.org/wp-content/uploads/2016/07/Memoria-proyecto-H%C3%A9rcules.pdf>. ↑ 49
- [26] Datos.gob.es, "Estudio de plataformas tecnológicas datos.gob.es". S. f. Ministerio de Industria, Turismo y Comercio. [En línea]. Disponible en: [http://datos.gob.es/sites/default/files/files/2\\_cms\\_01.pdf](http://datos.gob.es/sites/default/files/files/2_cms_01.pdf). ↑ 62
- [27] P.A. Gaona-García, C. E. Montenegro y H.W. González, "Hacia una propuesta de mecanismos para la autenticidad de objetos de aprendizaje en plataformas LCMS". *Revista Ingeniería*, vol. 19, no 1, p. 50-64. ↑ 63

---

### Jhon Francined Herrera Cubides

Ingeniero de Sistemas, Corporación Universitaria del Meta; especialista en Construcción de Software para Redes, Universidad Autónoma de Colombia; magíster en Ingeniería-Sistemas y Computación, Universidad de los Andes; estudiante del Doctorado en Ingeniería, Universidad Distrital Francisco José de Caldas; profesor asistente en el área de Programación, Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas; adscrito al grupo de investigación "GIIRA", donde realiza estudios sobre vinculación de datos. Correo electrónico: [jfherrera@udistrital.edu.co](mailto:jfherrera@udistrital.edu.co)

---

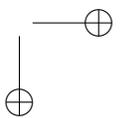
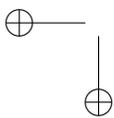
### Paulo Alonso Gaona-García

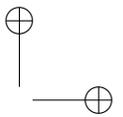
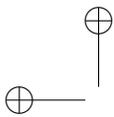
Ingeniero de Sistemas, Universidad Distrital Francisco José de Caldas; magíster en Ciencias de la Información y de las Comunicaciones, Universidad Distrital Francisco José de Caldas; profesor de tiempo completo adscrito a la Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas; doctor en Ingeniería de la Información y del Conocimiento, Universidad de Alcalá; sus áreas de interés se encuentran enfocadas en redes y comunicaciones, seguridad informática, e-learning, analítica visual de datos y Web semántica. Correo electrónico: [pagao-nag@udistrital.edu.co](mailto:pagao-nag@udistrital.edu.co)

---

### Kevin Gordillo Orjuela

Estudiante de Ingeniería de Sistemas adscrito al grupo de investigación GIIRA de la Universidad Distrital Francisco José de Caldas. Correo electrónico: [ksgordillo@correo.udistrital.edu.co](mailto:ksgordillo@correo.udistrital.edu.co)





Reporte de caso

## Análisis de Confiabilidad en Subestaciones Eléctricas Tipo Maniobra Implementando el Transformador de Tensión con Núcleo de Potencia

*Reliability Analysis in Maneuver Type Electrical Substations Implementing Tension Transformer with Power Core*

**Diego Viteri Toquica<sup>1</sup>, Cesar Garzón Bustos<sup>1</sup>, Alexander Narváez Cubillos<sup>1</sup>**

<sup>1</sup>Universidad Distrital Francisco José de Caldas. Bogotá - Colombia.

Correspondencia: deviterit@correo.udistrital.edu.co, cesagarzonb@correo.udistrital.edu.co, anarvaez@udistrital.edu.co

Recibido: 29-03-2016. Modificado: 06-07-2016. Aceptado: 12-10-2016

### Resumen

**Contexto:** Se pretende realizar un estudio técnico a un transformador de tensión con núcleo de potencia, usado para la alimentación de los servicios auxiliares de una subestación tipo maniobra, buscando mitigar la posibilidad de perder la continuidad del suministro de energía eléctrica y asimismo evitar la dependencia del agente distribuidor externo, el estudio se realiza entre la Empresa de Energía de Bogotá EEB y la Universidad Distrital Francisco José de Caldas.

**Método:** Se abordan dos etapas, primero un análisis de confiabilidad de la alimentación de los sistemas auxiliares de una subestación, realizado en tres escenarios: antes y después del uso del transformador y utilizando un grupo electrógeno exclusivo para los servicios auxiliares; la segunda etapa es una simulación en el software ATP para verificar el comportamiento del equipo en estado transitorio.

**Resultados:** La confiabilidad de la alimentación de los servicios auxiliares aumentó después de haber implementado el equipo, ya que la frecuencia de falla pasó de ser de 0,31[falla/año] con el sistema actual, a ser de 0,0025[falla/año].

**Conclusiones:** Implementar el transformador es conveniente para subestaciones que no cuenten con un sistema de alimentación de servicios auxiliares confiable ya que convierte a la subestación en auto dependiente, poner en funcionamiento este sistema en Colombia permitiría obtener un gran aumento en la confiabilidad de los servicios auxiliares de subestaciones tipo maniobra y, por tanto, una menor probabilidad de pérdida de carga.

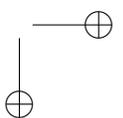
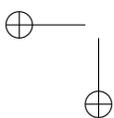
**Palabras clave:** Confiabilidad, Montecarlo, servicios auxiliares (SSAA), subestaciones eléctricas, transformador de tensión con núcleo de potencia (TTNP), transitorios.

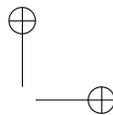
**Idioma:** Español



Citación: D. Viteri, C. Garzón, A. Narváez, "Análisis de confiabilidad en subestaciones eléctricas tipo maniobra implementando el transformador de tensión con núcleo de potencia" INGENIERÍA, vol. 22, no. 1, pp. 65-82, 2017.

©Los autores; titular de derechos de reproducción Universidad Distrital Francisco José de Caldas. En línea DOI: <http://dx.doi.org/10.14483/udistrital.jour.revving.2017.1.a04>





### Abstract

**Context:** It is intended to conduct a technical study of a voltage transformer power core, used for the supply of the auxiliary services of a maneuver type substation, hoping to mitigate the possibility of losing the continuous electricity power supply and likewise reducing the dependency with the external supplier. This is a result of a collaborative project between the Bogota Energy Company and the Universidad Distrital Francisco Jose de Caldas.

**Method:** The study consists of two stages, the first one is an analysis of reliability of the substation auxiliary systems, carried out in three scenarios: before and after use of the transformer and using a unique generator for auxiliary services. The second one is simulation using the ATP software to verify the behavior of the equipment in transient state.

**Results:** The reliability of auxiliary services increased after the equipment has been installed, as the failure rate went from 0.31 [failure/year] with the current system, to become 0.0025 [failure/year].

**Conclusions:** The implementation of transformer is suitable for substations that do not have a reliable power system auxiliary services as it makes the substation auto dependent. Implement this system in Colombia would get a big increase in the reliability of the auxiliary services in maneuver type substations and therefore a lower probability of loss.

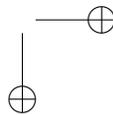
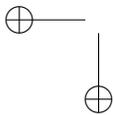
**Keywords:** Auxiliary services (SSAA), electrical substations, Monte Carlo, reliability, transient, voltage transformer core power (TTNP).

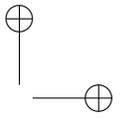
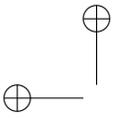
## 1. Introducción

Una subestación eléctrica representa un nodo dentro de un sistema eléctrico de potencia, en el que se puede transformar la energía eléctrica a niveles de tensiones apropiados para su transporte, distribución y consumo final. Igualmente se sabe que los SSAA (servicios auxiliares) son parte fundamental para el esquema y el correcto funcionamiento de la subestación por tal motivo son importantes para la confiabilidad de la misma; como caso de estudio en este documento se analizará a la subestación eléctrica de alta tensión que denominaremos “Sub\_A”, en la cual los SSAA, están siendo energizados por un generador eléctrico diésel de media tensión y la alimentación se lleva a la subestación por medio de una línea de transmisión. Estos dos elementos presentan los siguientes problemas:

- a. El generador eléctrico diésel es provisional, está previsto para dejar de funcionar eventualmente ya que pertenece a una empresa petrolera, que pretende retirar el equipo al término de sus labores en el área.
- b. La conexión del transformador de SSAA se realiza en la cola de un circuito rural de media tensión por lo cual la probabilidad de falla es bastante alta.
- c. Las torres por donde se llevó la línea de transmisión son propiedad de un agente distribuidor, por tal motivo este solicita un re-conector en la primera torre donde empieza la línea de SSAA para poder des-energizarla en el momento en que deban ejecutar algún trabajo de mantenimiento sobre sus torres o su línea.

Debido a estos problemas presentes en la Sub\_A, es necesario encontrar una solución que mitigue la posibilidad de perder la continuidad del suministro de energía eléctrica a los SSAA, y asimismo





evitar la dependencia del agente distribuidor externo. Es bajo este criterio que el **TTNP** (Transformador de tensión con núcleo de potencia) se convierte en una alternativa confiable técnicamente, a fin de solucionar la problemática presente en la Sub\_A. Es por eso, que se analizará la confiabilidad y el comportamiento del **TTNP**, como principal suministro de los SSAA de la Sub\_A.

### 1.1. Servicios auxiliares

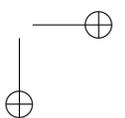
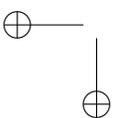
Los SSAA se definen como el conjunto de instalaciones y equipos que sirven para alimentar cargas en baja tensión de corriente alterna y continua, necesarias para la operación de la subestación. Se debe garantizar que en condiciones normales, de falla o mantenimiento, existan fuentes que alimenten las cargas que se consideren indispensables. También es importante resaltar que la confiabilidad de estos debe ser mayor a la de la subestación [1].

Las tensiones normalizadas para servicios auxiliares son:

- *Tensión en corriente continua* [2], [3]: 24V, 48V, 60V, 110V o 125V, 220V o 250V. (las más utilizadas en Colombia son 24V, 48V y 125V).
- *Tensión en corriente alterna* [2], [3]: 120/208V, 220/380V, 230/400V, 240/415V, 277/480V, 347/600V, (sistema trifásico tres o cuatro hilos).

Las formas tradicionales de alimentar los SSAA más comunes son las siguientes [4]:

- *Devanados terciario*: en ciertos casos los transformadores de potencia de la subestación son complementados con devanados terciarios en media tensión los que pueden usarse como fuente de alimentación para los SSAA. En algunos casos, no se recomienda usar este tipo de alimentación, puesto que se está dependiendo de las fluctuaciones mismas del sistema.
- *Transformador reductor*: normalmente esta alimentación se encuentra presente en subestaciones donde hay transformadores reductores para alimentar una subestación de distribución, por ejemplo 13,8kV. Uno de estos transformadores puede usarse como fuente de alimentación de los SSAA, aunque también son transformadores que dependen de las fluctuaciones del sistema y por tal motivo son poco recomendables.
- *Líneas aéreas de distribución urbanas o rurales*: las líneas de distribución de media tensión que pasen cerca de una subestación eléctrica de alta tensión, pueden ser utilizadas como fuente principal o suplementaria para alimentar los SSAA, se debe tener en cuenta que la confiabilidad de dicha línea de media tensión depende del agente distribuidor y de la distancia que exista desde el alimentador principal hasta el lugar donde será alimentado el transformador de SSAA de la subestación.
- *Grupo electrógeno*: los grupos electrógenos se usan como suministro de respaldo a otras fuentes de alimentación para garantizar el correcto funcionamiento del sistema. Sin embargo, en ocasiones poco frecuentes se puede dar el caso de requerir su utilización como fuente principal de alimentación para los SSAA.



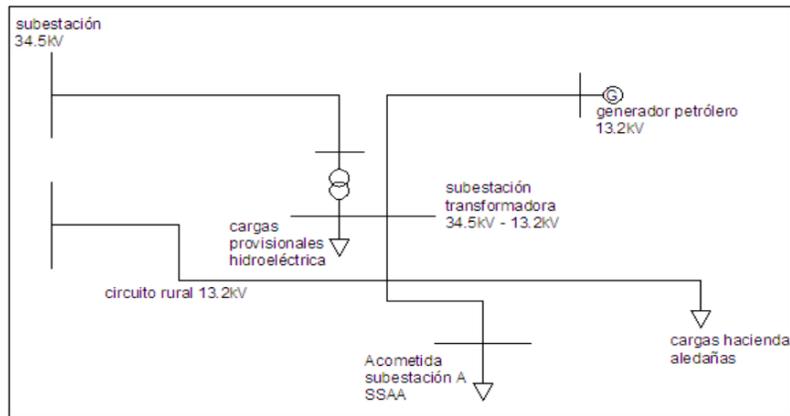


Figura 1. Alimentación de SSAA Sub\_A

Para el caso de estudio, la Sub\_A es una subestación tipo maniobra, por lo que no existe transformación a un nivel inferior al sistema de alta tensión. Este punto condiciona que las alimentaciones a los SSAA de C.A. se realicen desde sistemas eléctricos externos a la subestación. La solución que se adoptaría en este tipo de instalaciones es la alimentación desde una línea de media tensión con un alto grado de confiabilidad [5]. Dicha confiabilidad no se tiene en la Sub\_A, debido a que los SSAA están siendo alimentados por un generador eléctrico diésel que eventualmente será retirado. A su vez este generador alimenta una subestación tipo transformadora (34.5kV/13.2kV), que en uno de sus circuitos está alimentando provisionalmente la construcción de una central hidroeléctrica y en otro alimenta un circuito rural con líneas de 13.2kV, que finalmente alimentarán los SSAA de la Sub\_A, como se muestra en la Figura 1 [6]. Este circuito presenta varias desventajas, por lo que es necesario encontrar una fuente de alimentación confiable.

Internamente los SSAA de la Sub\_A son alimentados con un transformador de media/baja tensión “13.2kV/208-120V; 225kVA, que es energizado desde una línea de media tensión externa y que a su vez alimenta directamente 40 cargas, estas se distribuyen en 30 cargas no esenciales con un consumo de potencia eléctrica de 170KVA y 10 esenciales con un consumo de potencia eléctrica de 40KVA. Como respaldo a esta alimentación principal se cuenta con una planta diésel de 50kVA, 208/120V, que alimenta exclusivamente las cargas esenciales en caso de falla o mantenimiento del sistema principal de alimentación de SSAA.

La configuración de los SSAA en la Sub\_A, de barras acopladas con un solo alimentador de media tensión, que se define como dos barrajes unidos por un interruptor de acople. Esta configuración divide las cargas en esenciales y no esenciales, uno de los barrajes es energizado por la alimentación principal y de respaldo, en este, estarán ubicadas las cargas esenciales, el otro barraje tiene las cargas no esenciales y está energizado solamente por la alimentación principal [4]. En la Figura 2 se muestra esta configuración.

A continuación se explicará el funcionamiento, aplicaciones y principales características del TTNP, el cual en Colombia aún no se ha implementado.

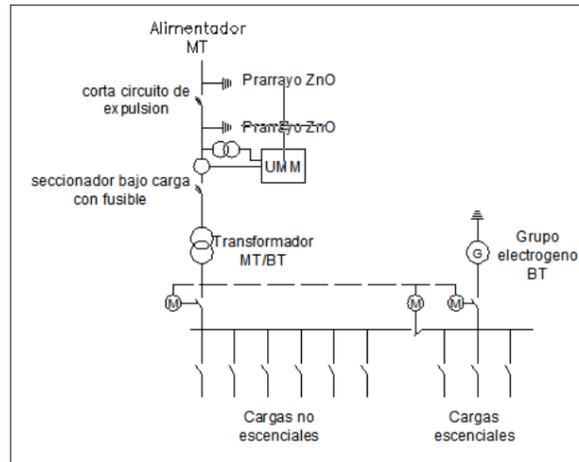


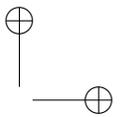
Figura 2. Configuración de barras acopladas con un solo alimentador de media tensión.

## 1.2. Transformador de tensión con núcleo de potencia

El **TTNP** es un transformador de voltaje monofásico que es capaz de suministrar potencia de hasta 333 kVA a un circuito en baja o media tensión directamente de un circuito de alta tensión, combina las características de un transformador de tensión y la capacidad de potencia de un transformador de distribución. Esto permite una conexión directa y confiable desde alta tensión (46kV–362kV), para obtener un suministro de energía con los siguientes rangos de tensión secundaria: 125/250V, 277V, 480V, o cualquier salida de hasta 600 V para baja tensión, y niveles normalizados de media tensión 11.4kV; 13.2kV o el nivel que sea necesario de acuerdo a la norma del lugar donde vaya a ser instalado para transmisión de la energía, se encuentra conectado línea a tierra y puede ser usado como una unidad individual para alimentar cargas monofásicas, o como banco trifásico capaz de suministrar hasta 1MVA de potencia eléctrica para grandes cargas [7].

El uso de los **TTNP** se remonta a finales de los 90, siendo una de sus aplicaciones más comunes, la electrificación rural para comunidades aisladas donde no es viable económicamente construir una subestación convencional para estos usuarios. Sin embargo muchas comunidades están ubicadas cerca de líneas de transmisión de alta tensión pero aun así no cuentan con suministro eléctrico en sus hogares. En estos casos la mejor solución fue el uso del **TTNP** [8], el cual se conecta directamente a las líneas de alta tensión y en su devanado secundario se obtienen tensiones normalizadas en rangos de media tensión para ser transportada a estas viviendas, ya que las características del transformador permite que en media tensión la energía eléctrica pueda ser transmitida a distancias de hasta 30km, para que allí sea transformada a baja tensión y finalmente sea consumida, casos de esta aplicación se encuentran en algunas comunidades en países de África Subsahariana como el Congo, donde se electrificaron siete comunidades con hospitales, escuelas, centros comunitarios por medio de una línea de transmisión de 245kV [9].

Sin embargo el **TTNP** tiene más aplicaciones de las cuales se pueden destacar:



- La alimentación de torres remotas de celular [10].
- Alimentación de SSAA en subestaciones donde no se tienen al menos dos fuentes confiables para el suministro de energía eléctrica de los SSAA [11] puede ser conectado directamente a los barrajes de la subestación y cumplir con la función de un transformador de tensión pero además suministrar varios kVA de potencia, debido a su núcleo.

El **TTNP** se perfila como una solución viable a la alimentación de los SSAA en la Sub\_A, ya que el diseño del mismo permite una ubicación conveniente dentro de la subestación para realizar el montaje. Además el equipo es suministrado por el fabricante con un devanado de potencia y uno o dos devanados de medición. Adicionalmente se pueden incluir algunas opciones tales como:

- Salidas de voltaje especiales en baja tensión
- Conectores especiales en baja tensión

El **TTNP** estándar viene con dos devanados secundarios de potencia, aislados que pueden ser conectados en serie o paralelo según sea el requerimiento de tensión a la salida, adicionalmente se puede entregar con una o dos bobinas de medición extraíbles con aislamiento entre estas y las bobinas de potencia para evitar fluctuaciones de tensión en las mediciones. Las opciones de conexión de cada devanado (potencia y medida) se muestran a continuación, Figura 3 y Figura 4 [12].

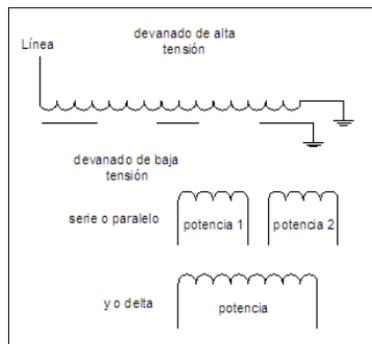


Figura 3. Conexión bobinas de potencia. Adaptada de [7]

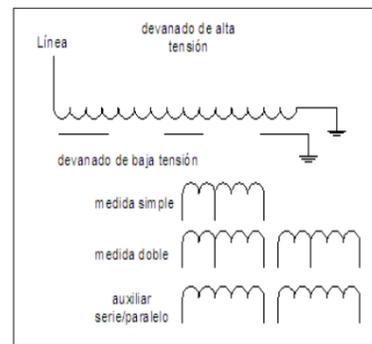
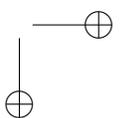
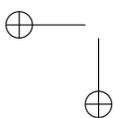


Figura 4. Conexión bobinas de medida. Adaptada de [7]

La selección del equipo se hace de acuerdo al BIL<sup>1</sup> de la subestación, gracias a la información brindada por la EEB, se conoce que el BIL de la Sub\_A corresponde a 1050 kV.

En general, en las subestaciones eléctricas, el **TTNP** es protegido por los descargadores de sobretensión y los cables de guarda, según varios fabricantes. Por lo que virtualmente en todas las aplicaciones, el primario se conecta directamente a la línea o en el caso de la Sub\_A al barraje [13].

<sup>1</sup>BIL: Nivel básico de aislamiento (Tensión soportada para impulso tipo rayo): es el valor pico de tensión soportada al impulso tipo rayo el cual caracteriza el aislamiento del equipo en lo que refiere a pruebas [15].



**Tabla I.** Nivel básico de aislamiento BIL [7]

		Tabla de datos del transformador									
BIL	sistema MSV	KV BIL	Estilo	10 KVA	25 KVA	50 KVA	100/125 KVA	160/167 KVA	200 KVA	250 KVA	315/333 KVA
1050	245 kV	1050 kV	Núcleo de potencia y demedida				•	•	•	•	•
							CI 0.5				

## 2. Método

A continuación se describe el proceso de las dos etapas del estudio realizado en la Sub\_A.

### 2.1. Evaluación de la confiabilidad de los SSAA en tres escenarios

La confiabilidad está relacionada con la disponibilidad de los equipos para brindar continuidad en el suministro de energía eléctrica cumpliendo con requerimientos y estándares de seguridad y calidad [14].

El propósito de este análisis es comparar la confiabilidad del sistema actual de alimentación de energía eléctrica de los servicios auxiliares de la Sub\_A, con otros dos posibles escenarios. El primero de estos es la implementación del **TTNP**; el segundo escenario se basa en la instalación de un grupo electrógeno propio de baja tensión. De esta manera se podrá determinar qué sistema presentaría la mayor confiabilidad.

Existen varios métodos para la estimación de la confiabilidad de un sistema eléctrico, para este caso de estudio se utilizaría el método de simulación de Montecarlo que consiste en generar números aleatorios de cualquier distribución de probabilidad o proceso estocástico para evaluar en forma numérica, indirecta o artificial un modelo matemático que permite estimar el comportamiento de un sistema o proceso que involucra variables estocásticas [16]. La base de este método es la generación de números aleatorios de cualquier distribución, que deben cumplir las siguientes propiedades:

- Uniformidad: pertenecen a una distribución de probabilidad uniforme definida entre 0 y 1.
- Independencia: los números generados no tienen relación entre sí.

Para iniciar con este análisis se buscaron reportes de salidas/restauración de cada uno de los componentes del sistema, en el caso de la Sub\_A se deben tener los datos de los equipos que componen el sistema utilizado para energizar los SSAA, al ser una subestación relativamente nueva, estos serán tomados de históricos de fallas en equipos similares, de la IEEE493 [17], se recomienda que los datos cubran un periodo mayor a un año, ya que, a mayor tiempo de los registros, mayor seguridad se obtendrá en los resultados de la simulación [18].

Para la simulación de Monte Carlo en este caso específico, la recolección de datos obedece a encontrar:

- Tasa de fallas  $\lambda$  [falla/año].
- Tasa de reparación  $\mu$  [reparación/año].

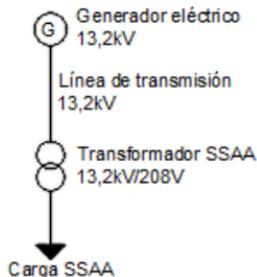


Figura 5. Sistema de alimentación SSAA convencional

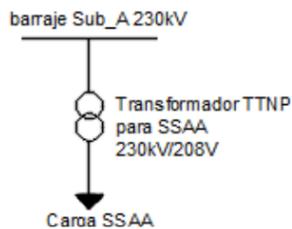


Figura 6. Alimentación SSAA con TTNP

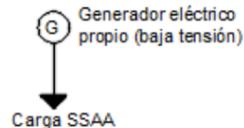


Figura 7. Alimentación SSAA con Grupo electrógeno propio.

Deben definirse de manera clara cada uno de los equipos que componen el sistema, en este caso de estudio se establecen tres escenarios. El primero, en el que los equipos son de media tensión y componen el circuito de abastecimiento de SSAA de la Sub\_A, (Figura 5). El segundo escenario el sistema cambia debido a que se contempla la instalación del TTNP (Figura 6). Para el último escenario analizado, se examina la instalación de un grupo electrógeno propio y de servicio continuo (Figura 7), como fuente principal para la alimentación de los servicios auxiliares en la Sub\_A.

Para el estudio de estos tres sistemas se contempla una simulación mediante el criterio de pérdidas de componentes  $n - 1$ , donde se considera que un elemento del sistema entra en estado de falla y permanece en el mismo por un tiempo de reparación, cumplido este tiempo el equipo recupera su disponibilidad y el sistema nuevamente se encuentra estable hasta la presencia de un nuevo estado de falla en cualquier elemento del sistema.

En la simulación de Montecarlo se permite incorporar modelos de confiabilidad de los elementos del sistema con cualquier número de estados. Estos modelos se definen por funciones de probabilidad para cada tiempo de transición entre estados. Sin embargo, en la mayoría de simulaciones reales se usan modelos de dos estados (servicio - falla) [19].

Para este caso de estudios se usó el modelo de dos estados (Figura 8), para cada uno de los equipos que componen cada sistema, ya que no existe información operativa ni registros para utilizar modelos más complejos.

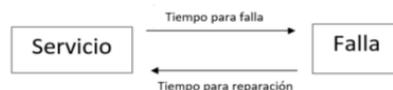


Figura 8. Modelo de dos estados para los equipos de la subestación

Cada componente tiene una secuencia operativa dentro del sistema, en esta se encuentran, el tiempo medio hasta fallas (MTTF o “médium time to failure”), el tiempo medio de reparaciones (MTTR o “médium time to repair”) y el tiempo medio entre fallas (MTBF o “médium time between failure”), como se definen en las ecuaciones (1), (2) y (3), respectivamente:

$$MTTF = \frac{1}{\lambda} \tag{1}$$

$$MTTR = \frac{1}{\mu} \tag{2}$$

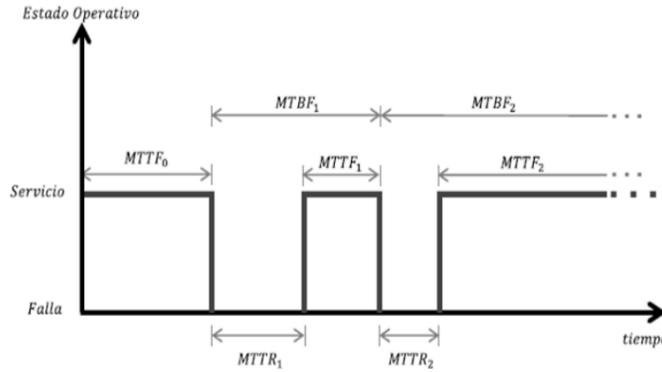


Figura 9. Estados operativos de los equipos o sistemas.

$$MTBF = MTTF + MTTR \tag{3}$$

En la Figura 9 [20], se ilustra el comportamiento de los equipos según estos índices.

Con los índices mencionados se llega al cálculo de la *Disponibilidad (A)* e *Indisponibilidad (U)* del sistema, antes y después de la instalación del equipo, las cuales vienen dadas por:

$$A = \frac{MTTF}{MTTF + MTTR} \tag{4}$$

$$U = \frac{MTTR}{MTTF + MTTR} \tag{5}$$

Una vez establecidos los índices se inicia con el proceso de simulación de los tres escenarios, el cual consistió de las siguientes mediciones:

- a. Encontrar las tasas de falla y reparación de cada elemento que conforma los dos sistemas de estudio (Tabla II, Tabla III y Tabla IV. )

Tabla II. Tasa de falla y reparación sistema 1

Elemento	$\lambda$ [Falla/Año]	$\mu$ [Rep./Año]	FUENTE
Línea	0,37362 <sup>a</sup>	826,41	[17]
Transformado	0,04450	4460,5	[21]
Generador diésel	0,01350	478	[17]

<sup>a</sup>Este dato es calculado para el caso de estudio, ya que el índice registrado es 0,1437[Falla/Año] por cada 1000 (ft) de línea

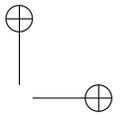
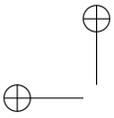
Tabla III. Tasa de falla y reparación sistema 2

Elemento	$\lambda$ [Falla/Año]	$\mu$ [Rep./Año]	FUENTE
Barra	0,0017958	730	[22]
Transformador con núcleo de potencia	0,0014	1752	[23]

Tabla IV. Tasa de falla y reparación sistema 3

Elemento	$\lambda$ [Falla/Año]	$\mu$ [Rep./Año]	FUENTE
Grupo electrógeno propio	0.1286	478	[17]

Los elementos anteriormente mencionados si bien no reflejan los datos exactos del caso de estudio, son consultas realizadas con equipos similares con los que se pretende semejar a los reales



en la Sub\_A. Cabe resaltar que de tener registros reales de cualquier subestación a la cual se le pretenda realizar un estudio de confiabilidad, es aconsejable usarlos.

**b. Cálculos de tiempos de falla y reparación para cada componente de los sistemas.**

Los números aleatorios, fundamentales en la simulación de Monte Carlo, son generados en Excel mediante la función aleatorio, la cual genera números aleatorios sin ningún patrón de semilla entre ellos, en un rango de 0-1.

Una vez generados estos números se analiza la variable aleatoria de interés teniendo en cuenta que toda función de distribución de probabilidad evaluada en cualquier valor de  $X$  es igual a un número que representa la probabilidad entre 0 y 1 [24], por lo que:

$$F(x) = \text{Aleatorio} \quad (6)$$

Por lo tanto, se puede obtener un tratamiento a la variable  $X$  hallando la función inversa de la distribución de probabilidad:

$$x = F^{-1}(\text{Aleatorio}) \quad (7)$$

Para este caso en particular se usó la función de distribución exponencial, ya que su comportamiento asemeja la vida útil de cada elemento de los sistemas mientras aumenta su tiempo de uso en relación a las fallas que presenta durante este tiempo, la función inversa de la distribución exponencial es:

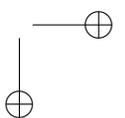
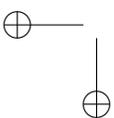
$$x = -\frac{1}{h} * \ln(\text{Aleatorio}) \quad (8)$$

Para generar los tiempos de falla y reparación aleatorios, en la fórmula de la distribución exponencial inversa se debe cambiar el parámetro “ $h$ ”, que es la tasa de eventos o el inverso del parámetro de escala, por los índices de falla ( $\lambda$ ) y reparación ( $\mu$ ) respectivamente encontrados para cada elemento de los sistemas [25]. Para este caso se decidió tomar un tiempo de expectativa de 50 Años o 438000 Horas en los sistemas planteados.

## **2.2. Simulación en ATP para determinar el comportamiento del TTNP ante una sobretensión por descarga atmosférica**

Las sobretensiones por descargas atmosféricas son las tensiones fase-fase o fase-tierra que son superiores al nivel de tensión de diseño del sistema de potencia ( $V_m$ ), son de origen externo, y son causadas por la caída de un rayo sobre o alrededor de un punto del sistema de potencia y se caracterizan por ser de muy corta duración [26].

Para realizar el análisis se debe tener en cuenta el motivo por el cual el rayo impacta directamente una línea de fase ya que según estudios realizados por la empresa interconexión eléctrica S.A (ISA) las descargas atmosféricas son el principal motivo para que una línea de alta tensión experimente una salida del sistema de potencia.



Entre las salidas producidas por descargas atmosféricas se deben destacar las más importantes que son las salidas producidas por fallas de apantallamiento y las debidas a flameo inverso [27]. Las producidas por fallas de apantallamiento son debidas a la caída de un rayo directamente en el conductor de fase por falla en el diseño del cable de apantallamiento, esto puede causar la salida inmediata de la línea ya que el rayo puede tener una corriente de 30 kA característicamente por lo que puede generar ondas de tensión a lado y lado del cable de fase y dependiendo de la impedancia característica del conductor se generará una sobretensión que puede superar los niveles de BIL característicos de una línea de 220kV, por tal razón se puede producir un flameo inverso entre conductores de fase lo cual acarrearía la salida inmediata de toda la línea [28].

Por otra parte, si la descarga atmosférica cae sobre el cable de guarda reflejara ondas a lado y lado que llegarán a la cruceta y tratarán de ser descargadas a tierra por medio de la bajante o de la misma torre, dependiendo de la impedancia de cada uno de estos elementos, estas ondas producidas por el rayo generarán tensiones que pueden ser bastante altas, si la tensión producida en la cruceta es mucho mayor que la tensión de las fases se produce el flameo inverso (back flashover), este flameo puede ocasionar dependiendo de la configuración de las protecciones la salida de la línea [27].

Para el estudio se utilizarán las sobretensiones provocadas por impacto directo sobre un conductor de fase (Figura 10) [28], ya que este sería el peor escenario para el sistema de potencia donde estaría instalado el TTNP y el objetivo es verificar cual será el comportamiento del devanado secundario del transformador debido a una descarga atmosférica en cualquier punto de dicho sistema.

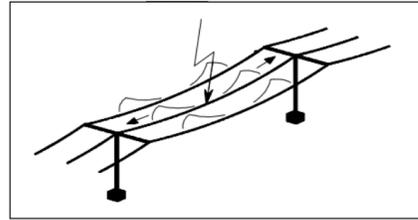


Figura 10. Impacto directo en un conductor de fase

Para la simulación de la sobretensión, se tendrán en cuenta posiciones de impacto aleatorias den-

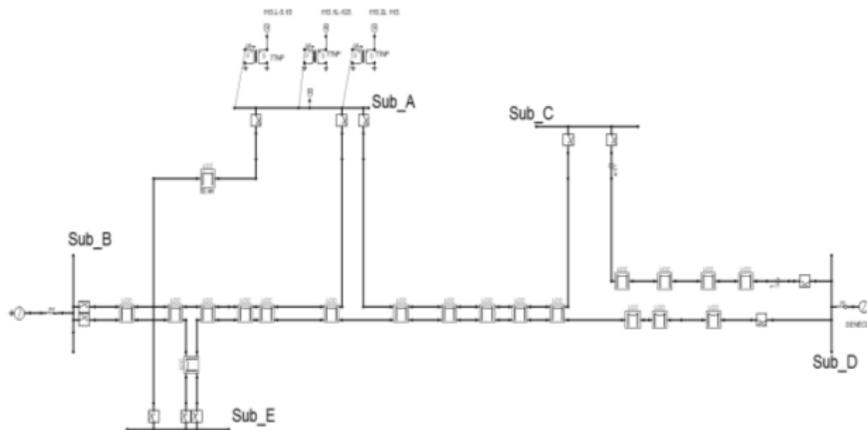


Figura 11. Sistema de potencia bajo estudio.

tro del sistema, donde se aplicará el impulso tipo rayo.

En la (Figura 11) se muestra el sistema de potencia estudiado para la simulación de las sobretensiones causadas por las descargas atmosféricas.

En la parte superior se observa la Sub\_A con los tres **TTNP** en cada una de sus fases, correctamente modelado de acuerdo a los datos de las pruebas eléctricas de corto circuito y circuito abierto realizadas por el fabricante del equipo.

Igualmente se modela un sistema de potencia real con sus respectivas líneas por tramos y longitudes de acuerdo a datos del sistema de la EEB, para poder obtener resultados bastante realistas.

### 3. Resultados

#### 3.1. Valoración de confiabilidad

Al simular los elementos de cada sistema en conjunto y con los tiempos generados aleatoriamente, se registra que elemento tiene el menor tiempo de falla, esto debido a que nuestro criterio de contingencias es  $n - 1$ , mediante este proceso se encuentran los equipos en falla, como se indica a continuación:

**Tabla V.** Elemento en falla bajo criterio  $n - 1$  del sistema 1

Tiempo aleatorio de falla [Falla/Hora]		Mínimo	Equipo en	Tiempo aleatorio	Tiempo entre	
Línea	Transformador	Generador diesel [Falla/Hora]	falla	reparación [Rep./Hora]	fallas [Hora]	
4458,62	85246,50	18642,77	4458,62	línea	19,44	4478,07

**Tabla VI.** Elemento en falla bajo criterio  $n - 1$  del sistema 2

Tiempo aleatorio de falla [Falla/Hora]		Mínimo	Equipo en	Tiempo de reparación	Tiempo entre
Barra	Transformador de tensión con núcleo de potencia	[Falla/Hora]	falla	aleatorio [Rep./Hora]	fallas [Hora]
22345464,65	4220167,24	4220167,24	SSVT	5,63	4220172,88

**Tabla VII.** Elemento en falla bajo criterio  $n - 1$  del sistema 3

Tiempo aleatorio de falla [Falla/Hora]		Mínimo	Equipo	Tiempo de reparación	Tiempo entre
Generador	Servicio continuo	[Falla/Hora]	en falla	aleatorio [Rep./Hora]	fallas [Hora]
36319,66	36319,66	36319,66	Generador S.C.	34,14	36353,80

Como se puede observar en la Tabla V , Tabla VI y en la Tabla VII , el programa en Excel selecciona el elemento con menor tiempo de falla e identifica el elemento fallado en cada sistema.

El proceso se repite con los tiempos de reparación generados aleatoriamente, ya que estos se sumarán al tiempo de falla del elemento mencionado por el programa, esto con el objetivo de establecer un tiempo de recuperación del sistema o tiempo entre fallas.

Una vez se establece el tiempo entre fallas de cada sistema se realizan las iteraciones tantas veces sean necesarias para alcanzar el tiempo de estudio (50 Años o 438000 Horas), cabe resaltar que todos los índices se convirtieron en horas.

Durante el análisis se llegó a establecer que, para cumplir el tiempo de estudio cercano al deseado, en el Sistema 1 fueron necesarias veinte iteraciones; estableciendo esto para el Sistema 2, se realizaron el mismo número de iteraciones, pero se comprobó que en la primera iteración para este sistema se sobrepasó el tiempo de estudio deseado puesto que los elementos de este tienen tasas de falla más bajas en comparación con el Sistema 1.

Con los tiempos generados en cada sistema se calcularon los MTTF, MTTR y MTBF (Tabla VIII, Tabla IX y Tabla X), con estos y mediante la aplicación de las formulas establecidas anteriormente las aproximaciones a las tasas de falla y restauración del sistema entero.

Los índices LOLE y LOLP, corresponden a índices de confiabilidad del suministro de energía [29].

**LOLP (Loss Of Load Probability):** probabilidad de que la carga del sistema exceda la capacidad de suministro del mismo. No se da cobertura completa a la demanda de suministro.

**LOLE (Loss Of Load Expectation):** número de días en el cual el pico de la demanda se espera que supere la capacidad de suministro disponible.

**Tabla VIII. Resultados del sistema 1**

MTTF [Año/Falla]	3,18
$\lambda$ [Falla/Año]	0,31
MTTR [Año/Rep.]	0,0000018
$\mu$ [Rep./Año]	556352,8
Disponibilidad	0,99
Indisponibilidad	0,00040
Frecuencia de falla	0,31
LOLP	0,00039
LOLE	3,47

**Tabla IX. Resultados del sistema 2**

MTTF [Año/Falla]	392,62
$\lambda$ [Falla/Año]	0,0025
MTTR [Año/Rep.]	1,45E-08
$\mu$ [Rep./Año]	68788393,7
Disponibilidad	0,99
Indisponibilidad	2,75E-06
Frecuencia de falla	0,0025
LOLP	2,75E-06
LOLE	0,024

**Tabla X. Resultados del sistema 3**

MTTF [Año/Falla]	6,43
$\lambda$ [Falla/Año]	0,15
MTTR [Año/Rep.]	8,86E-07
$\mu$ [Rep./Año]	1127707,91
Disponibilidad	0,99
Indisponibilidad	0,00021
Frecuencia de falla	0,12
LOLP	0,000179
LOLE	1,56

Para un caso óptimo se espera que en los sistemas de estudio se obtengan valores bajos para LOLE y LOLP, de los tres escenarios estudiados el Sistema 2 resultó ser el óptimo, el Sistema 3, también presenta unos mejores índices en comparación al Sistema 1, pero no tanto como el Sistema 2. Además es válido afirmar que el grupo electrógeno durante su funcionamiento desprende sustancias tóxicas como: óxidos de nitrógeno, hollín, monóxido de carbono, hidrocarburos, compuestos de azufre y plomo [30], por lo que no solo, no es la opción más confiable como alimentación principal de los SSAA, sino que también es bastante contaminante, por lo tanto se puede concluir que la confiabilidad de la alimentación de los SSAA de la Sub\_A, mejoraría al usar el *transformador de tensión con núcleo de potencia*.

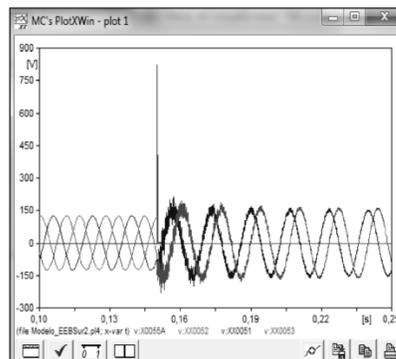
### 3.2. Simulación para sobretensiones tipo rayo

A continuación se analizarán los resultados obtenidos aplicando el impulso de tensión tipo rayo, estos resultados son exclusivamente del devanado secundario del equipo bajo estudio, debido a que

es este devanado el que estará finalmente energizando las cargas de SSAA de la Sub\_A.

- **Descarga atmosférica sobre la barra de la Sub\_A**

Se observa que el devanado secundario de los **TTNP** inicialmente están trabajando sin ningún problema, al ocurrir una descarga atmosférica sobre la barra de la sub\_A se observa que de manera inmediata la fase sobre la cual ocurre la descarga llega a valores aproximados a cero mientras que las otras dos fases sufren una fluctuación elevadas debido a la interacción del rayo por flameo sobre las otras fases (Figura 12), después de que se disipa el rayo se ve cómo las fases tienden a estabilizarse pero se mantienen en valores elevados de tensión debido a que el sistema trata de compensar la pérdida de la fase sobre la cual cae la descarga atmosférica. En la Tabla XI se pueden ver reflejados los valores de tensión en los tiempos en los cuales se deben ver los picos de tensión de cada fase y se observa claramente cómo las dos fases que no sufrieron el impacto elevan su nivel de tensión a valores alrededor de los 150V.



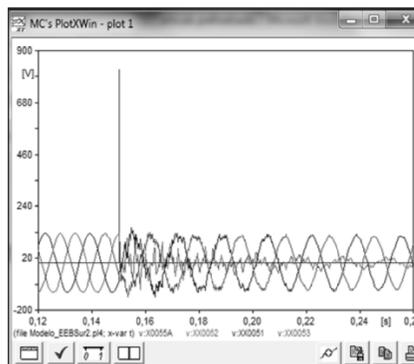
**Figura 12.** Sobretensión tipo rayo sobre la barra de la sub\_A fase A

**Tabla XI.** Tensión de fases para falla en fase A barra Sub\_A

FASES	$\pi/3$	$2\pi/3$	$\pi$
A	-4,17E-14V	-4,17E-14 V	-8,44E-11 V
B	145,36 V	-54,52 V	-117,14 V
C	-26,6 V	156,8 V	-123,23 V

- **Descarga atmosférica sobre la barra de la Sub\_B a 82 km de Sub\_A**

En este caso se observa el comportamiento del **TTNP**, cuando ocurre una descarga atmosférica sobre una de sus fases. En la fase donde se presenta la descarga atmosférica se evidencia una fluctuación, además se observa en la Tabla XII y según la (Figura 13), la tensión de la fase impactada es de 55,49V máximo y con el paso del tiempo se va estabilizando, esto debido a los efectos del rayo sobre las otras fases y al ser este un nodo de generación se observa cómo hay fluctuación en tres periodos de la onda y a partir de ahí esta empieza a atenuar en unos valores de tensión que están por debajo de la tensión nominal de dichas fases.



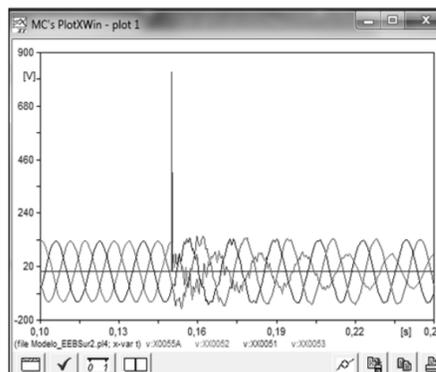
**Figura 13.** Descarga atmosférica en la Sub\_B

**Tabla XII.** Tensión de fases por descarga atmosférica en Sub\_B fase A

FASES	$\pi/3$	$2\pi/3$	$\pi$
A	55,49 V	0,18 V	-6,88 V
B	84,6 V	-91,5 V	-35,6 V
C	-85,2 V	114,11 V	-31,36 V

- **Descarga atmosférica sobre la barra de la Sub\_B a 82Km de Sub\_A**

Para la descarga que ocurre entre la Sub\_A y la Sub\_C a 142.1 Km de la Sub\_A (Figura 14) se observa la afectación sobre la fase en la cual impacta directamente la descarga y las pequeñas fluctuaciones sobre las otras fases debido a los efectos del rayo, luego de algunos milisegundos después de ocurrir el impacto, el sistema persiste en mantener una tensión sobre la fase que impacta el rayo, mientras las otras fases aumentan su tensión para tratar de suplir la pérdida que está generando la fase en falla, como se puede observar tanto en la Figura 14. En la Tabla XIII, es evidente que por la distancia en la que impacta el rayo y en comparación con las otras simulaciones la fase en falla no pierde totalmente su tensión y por lo tanto las otras fase no se ven forzadas a elevar su nivel de tensión, tanto como si lo tuvieron que hacer cuando impacto el rayo directamente en la barra de la Sub\_A y cuando impacto muy próximo a la misma subestación.



**Figura 14.** Comportamiento del TTNP ante una descarga atmosférica a 142.1km de la Sub\_A.

**Tabla XIII.** Tensión de fases salida por descarga atmosférica en fase A a 142km de Sub\_A.

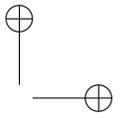
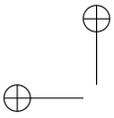
FASES	$\pi/3$	$2\pi/3$	$\pi$
A	-59,60V	-17,28V	59,85V
B	121,2V	-6,9V	-68,57V
C	-54,78V	145,76V	60,76V

## 4. Conclusiones

Según el estudio de confiabilidad realizado, el TTNP cumple con las exigencias de diseño, mantenimiento y durabilidad con respecto a los sistemas de alimentación de SSAA actuales en la Sub\_A. También se demuestra que implementando el TTNP la probabilidad de pérdida de carga es menor en comparación a los demás escenarios, por lo tanto la alimentación de SSAA será más confiable. Es recomendable aplicar el estudio con datos reales de cualquier subestación en la que se pretenda implementar este sistema, como ya se había mencionado anteriormente, ya que de acuerdo a la realidad y veracidad de los registros ingresados, los resultados que se obtendrán serán más concretos.

Si bien se demuestra que el TTNP es la mejor opción técnica para mejorar la confiabilidad de los servicios auxiliares, también se recomienda realizar un estudio de factibilidad económica que compare los tres escenarios que se estudiaron, para determinar definitivamente si el TTNP, es la opción más rentable, para alimentar los servicios auxiliares de una subestación tipo maniobra.

Para el caso de la Sub\_A es recomendable implementar el TTNP, ya que este elimina la dependencia de agentes distribuidores en media tensión, convirtiendo la subestación en auto dependiente, pues la alimentación de los SSAA, es directamente tomada de la barra de la misma, simplificando el circuito de alimentación de los mismos sin generar contaminación, como sí lo haría implementar un grupo electrógeno.



Dependiendo de la magnitud de la corriente del rayo y de factores de diseño de la línea como lo serían distancias de seguridad, diámetro de los conductores, impedancia característica de cada uno de los componentes del sistema asimismo se verá afectado ante una descarga atmosférica y de esto dependerá la gravedad de la falla y de si la línea sea afectada parcial o totalmente.

Las descargas atmosféricas rara vez tendrán impacto directo sobre una barra en una subestación debido al diseño de apantallamiento, por lo cual las simulaciones en las que se utilizó el impacto directo en las barras tienen una probabilidad casi nula de suceder.

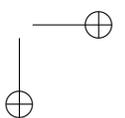
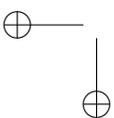
Es preciso recordar que el TTNP, tiene aplicaciones que son recomendables estudiar posteriormente, una de ellas es la electrificación rural, la cual en nuestro país tendría un gran impacto social en las zonas no interconectadas, desde este punto de vista es imprescindible saber que para una adecuada educación energética se deben contar con nuevas tecnologías que conjuguen las necesidades del sector energético y los posibles nuevos escenarios que permitan atender las problemáticas ambientales generadas por antiguos modelos de energización [31]. También destacando que la tecnología sería nueva en el país se recomienda realizar el estudio técnico para la utilización en alimentación de antenas de telecomunicaciones con el fin de retirar el cableado en las redes de telecomunicaciones y utilizar antenas de transmisión de señal inalámbrica.

### Reconocimientos

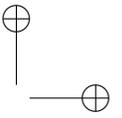
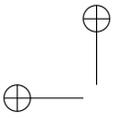
Se da formal reconocimiento a la Universidad Distrital Francisco José De Caldas y a la Empresa de Energía de Bogotá, en especial al Ingeniero Edgar Yesid Torres por su disposición, compromiso y apoyo, durante la elaboración de este artículo.

## Referencias

- [1] Martínez, C. *Plan de mantenimiento de servicios auxiliares de las subestaciones del SEN*. Tesis Pregrado, Universidad Simón Bolívar, Sartenejas, 2012. ↑67
- [2] IEC 60694, *Common specifications for high-voltage switchgear and controlgear standards*. Geneva, 1996. ↑67
- [3] ANSI C84.1, *For Electric Power Systems and Equipment-Voltage Ratings (60 Hertz)*. Rosslyn, VA, 2011. ↑67
- [4] M. Villegas S.A., *Subestaciones de alta y extra alta tensión (Segunda edición)*. Medellín, Impresiones Gráficas Ltda., 2003. ↑67, 68
- [5] EDP HC Energía, “Servicios Auxiliares de Subestaciones ET/4001”. S.f. [En línea]. Disponible en: <http://www.edphcenergia.es/recursosedp/doc/distribucion-luz/20130917/subestaciones-y-centros-d/servicios-auxiliares-de-subestaciones.pdf> ↑68
- [6] EEB, Gerencia de mantenimiento, *Visita a la zona SUR*, Empresa Energía de Bogotá, Bogotá, D.C., 2015. ↑68
- [7] ABB Ltd, “Types SSVT and SSMV oil-filled station service voltage transformers Auxiliary power from high voltage transmission lines”. S.f. [En línea]. Disponible en: [http://www02.abb.com/global/brabb/brabb155.nsf/bf177942f19f4a98c1257148003b7a0a/64d9fd97c7e797bf83257ebb00639c19/\\$FILE/Cat%20A1logo+-+SSVT.pdf](http://www02.abb.com/global/brabb/brabb155.nsf/bf177942f19f4a98c1257148003b7a0a/64d9fd97c7e797bf83257ebb00639c19/$FILE/Cat%20A1logo+-+SSVT.pdf) ↑69, 70, 71
- [8] Juma, M., and Trevor, C., “Implication of Using Auxiliary Service Voltage Transformer Sub-Stations for Rural Electrification”. *International Journal of Energy and Power Engineering*, vol. 4, pp. 1-11, Mar. 2015. ↑69
- [9] Artech, “Calálogo Transformadores de tensión para servicios auxiliares”. 2015. <http://artech.com/es/productos/transformadores-de-tensi%20B3n-para-servicios-auxiliares> ↑69
- [10] ABB Ltd., “ SSVT and SSMV product catalog oil-filled station service voltage transformers Auxiliary power from high voltage transmission lines”. S. f. [En línea]. Disponible en: [https://library.e.abb.com/public/8b3c5fb893624902a5d4a0cb7f2ace5e/SSVT%20Family%20Brochure\\_PPHVITSSV1014BR\\_final.pdf](https://library.e.abb.com/public/8b3c5fb893624902a5d4a0cb7f2ace5e/SSVT%20Family%20Brochure_PPHVITSSV1014BR_final.pdf) ↑70



- [11] Henao, W., “Transformadores de Voltaje de Alta Tensión para suministro directo de Potencia en Baja y Media Tensión”. [En línea]. Disponible en: <http://new.abb.com/docs/librariesprovider78/chile-documentos/novenas-jornadas-tecnicas-2015/pp/william-henao.pdf?sfvrsn=2> ↑70
- [12] ABB Ltd., “Station Service Voltage Transformer (SSVT) general protection”. S. f. [En línea]. Disponible en: [https://library.e.abb.com/public/226de1d6648b678e85257c71004b2b3a/SSVT%20General%20Protection\\_PPHVI\\_TSSV0312TD.pdf](https://library.e.abb.com/public/226de1d6648b678e85257c71004b2b3a/SSVT%20General%20Protection_PPHVI_TSSV0312TD.pdf) ↑70
- [13] ABB Group, *SSVT- Transformador de Voltaje para potencia auxiliar*. Jornadas técnicas de ABB, Lima, Perú, 2015. ↑70
- [14] Zapata, C., “Conceptos generales” in *Confiabilidad Sistemas Eléctricos de potencia*, 1st ed., Pereira, Universidad Tecnológica de Pereira, 2011, pp. 1-8. ↑71
- [15] Ruiz, D., y Lozano, D., *Coordinación de aislamiento*. 2007. [En línea]. Disponible en: <http://cursa.ihmc.us/rid=1L7681P9H-2BT34NJ-1MP3/COORDINACION-DE-AISLAMIENTO.pdf> ↑70
- [16] Faulín, J., y Ángel, J., “Simulación de Montecarlo [http://www.uoc.edu/in3/emath/docs/Simulacion\\_MC.pdf](http://www.uoc.edu/in3/emath/docs/Simulacion_MC.pdf) ↑71
- [17] IEEE, “Summary of equipment reliability data” in IEEE Std 493-1997 (Revision of IEEE Std 493-1990), pp. 37-78, Dec. 16, 1997. ↑71, 73
- [18] Zapata, C., Garces, L., y Gómez, O., “Modelamiento de componentes de sistemas compuestos generación - transmisión para estudios de confiabilidad”. *Scientia et Technica*, 38(25), pp. 53-58, 2004. [En línea]. Disponible en: [http://www.utp.edu.co/planeamiento/prod\\_aca/articulos/artc2Lina.pdf](http://www.utp.edu.co/planeamiento/prod_aca/articulos/artc2Lina.pdf) ↑71
- [19] Zapata, C., y Gómez, V. “valoración de confiabilidad de subestaciones eléctricas utilizando simulación de montecarlo”. *Scientia et Technica* 3(32), pp. 67-72. 2006. [En línea]. Disponible en: <http://revistas.utp.edu.co/index.php/revistaciencia/article/view/6205/3595> ↑72
- [20] Ruiz, J., *Propuesta metodológica para el cálculo y gestión de la confiabilidad del suministro energético en instalaciones industriales*. Tesis Especialización, Universidad Pontificia Bolivariana, Medellín, 2013. ↑73
- [21] Zapata, C., Cataño, D. y Suárez, H., “Índices de confiabilidad de transformadores de distribución”. *Mundo Eléctrico*, 57(57), pp. 102-106, 2004. [En línea]. Disponible en: <http://academia.utp.edu.co/planeamiento/files/2014/01/IndicesConfiabilidadTrafos.pdf> ↑73
- [22] Muñoz, G., *Análisis de confiabilidad de arreglos de barras e interruptores en extra alta tensión mediante árboles de falla*. Tesis Pregrado, Universidad de Chile, Santiago de Chile, 2012. ↑73
- [23] Martínez, P., *Mejoras en el cálculo de índices de fiabilidad en redes malladas de distribución de energía eléctrica*. Tesis doctoral, Universidad de Sevilla, Sevilla, 2013. ↑73
- [24] Zapata, C., *Análisis probabilístico y simulación*. Pereira, Universidad Tecnológica de Pereira, 2010, pp. 38-109. ↑74
- [25] Billinton, R., and Allan, R., *Reliability Evaluation of Power Systems (2nd ed.)* New York and London, Plenum Press, 1996, pp. 400-417. ↑74
- [26] Blandón, J. “Coordinación de aislamiento norma iec 60071-2 vs ingeniería por desempeño”. S. f. [En línea]. Disponible en: [http://ingenieria.udea.edu.co/altae2009/documentos/conferencias\\_magistrales/conferencia\\_ingenieria\\_desempeno\\_jaime\\_blandon.pdf](http://ingenieria.udea.edu.co/altae2009/documentos/conferencias_magistrales/conferencia_ingenieria_desempeno_jaime_blandon.pdf) ↑74
- [27] Mejía, A. *Características de las descargas atmosféricas y su efecto sobre las líneas*. 2005. [En línea]. Disponible en: <http://www.gamma.co/caracteristicas-las-descargas-atmosfericas-efecto-las-lineas-transmision/> ↑75
- [28] González, F. *Evaluación estadística del comportamiento de líneas aéreas de distribución frente a sobretensiones de origen externo*. Tesis doctoral, Universitat Politècnica de Catalunya, Cataluña, 2001. ↑75
- [29] Ventosa, M., y Ramos, A. “Modelos de sistemas de energía eléctrica”. 2006. [En <http://www.iit.comillas.edu/aramos/presentaciones/Tema%20M5%20Indices%20de%20fiabilidad.pdf> ↑77
- [30] Martínez, M., Fernández, A., Molina, E., y García, R., “Grupos electrógenos y su impacto ambiental”. *Higiene y Sanidad Ambiental* 7, pp. 217-221, 2007. [En línea]. Disponible en: [http://www.salud-publica.es/secciones/revista/revistaspdf/bc51015c2c43ce3\\_Hig.Sanid.Ambient.7.217-221\(2007\).pdf](http://www.salud-publica.es/secciones/revista/revistaspdf/bc51015c2c43ce3_Hig.Sanid.Ambient.7.217-221(2007).pdf) ↑77
- [31] Gallego, A. y Castro, J. “Sobre el rol innovador de la educación energética para la investigación en ingeniería”. *Ingeniería*, Vol. 19, No. 2, pp. 147-163, 2014. ↑80



---

### **Diego Ernesto Viteri Toquica**

Estudiante de Ingeniería eléctrica de la Universidad Distrital Francisco José de Caldas; se desempeñó como practicante Universitario en la Empresa de Energía de Bogotá donde desarrolló protocolos de mantenimiento para equipos de patio de alta tensión. Actualmente es ingeniero en el área de diseño en la Empresa ACJ High Voltage, Tenjo Cundinamarca, Colombia.

Correo electrónico: deviterit@correo.udistrital.edu.co

---

### **César Andres Garzón Bustos**

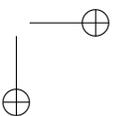
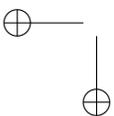
Estudiante de Ingeniería Eléctrica de la Universidad Distrital Francisco José de Caldas; se desempeñó como ingeniero residente en empresas de ingeniería eléctrica.

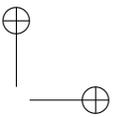
Correo electrónico: cesagarzonb@correo.udistrital.edu.co

---

### **Eider Alexander Narváez Cubillos**

Ingeniero Electricista, Universidad Nacional de Colombia; magíster en Ingeniería Eléctrica, Universidad los Andes; actualmente adelanta sus estudios de doctorado en la Universidad Nacional de Colombia en el tema de sistemas de almacenamiento de energía y se encuentra vinculado al grupo de Compatibilidad Electromagnética de la misma Universidad. Ha sido profesor en diferentes Universidades en la ciudad de Bogotá y actualmente es profesor de planta en la Universidad Distrital Francisco José de Caldas, de Bogotá, donde ha trabajado en las áreas de máquinas eléctricas, instalaciones eléctricas y sistemas de potencia. Perteneció al grupo de Investigación de LIFAE de la misma institución. Correo electrónico: anarvaez@udistrital.edu.co





Investigación

**Monitoreo del Clima Espacial desde Colombia Mediante Radio Receptores Butterworth de Orden Superior**  
*Space Weather Monitoring from Colombia Through Higher Order Butterworth Radio-Receivers*

**Andrés Felipe Jaramillo Alvarado<sup>1</sup>, Hamilton David Galvis Rodríguez<sup>1</sup>, Edwin Andrés Quintero Salazar\*<sup>1</sup>**

<sup>1</sup> Grupo de Investigación en astroingeniería alfa orión, Observatorio Astronómico, Universidad Tecnológica de Pereira.

\*Correspondence: equintero@utp.edu.co

Recibido: 25-07-2016. Modificado: 15-09-2016. Aceptado: 12-10-2016.

Resumen

**Contexto:** Las fulguraciones solares y las eyecciones de masa coronal liberan al espacio grandes cantidades de radiación que, al alcanzar las capas exteriores de la atmósfera terrestre, alteran sus características eléctricas. El conjunto de fenómenos derivados de esta interacción se conoce como clima espacial y sus consecuencias incluyen fallas en los sistemas de geolocalización, las telecomunicaciones, las operaciones vía satélite, el seguimiento espacial, la radionavegación y la sobrecarga de redes eléctricas.

**Método:** En este trabajo se presenta el desarrollo de un radio receptor que permite adelantar el monitoreo del clima espacial desde Colombia. Además, se describe una metodología novedosa para el diseño de filtros analógicos de orden superior de característica Butterworth, a partir de la interconexión en paralelo de bancos de filtros de primero y segundo orden, obteniéndose la función de transferencia deseada.

**Resultados:** El radio receptor desarrollado se instaló en la estación de monitoreo del clima espacial del Observatorio Astronómico de la Universidad Tecnológica de Pereira (OAUTP), Colombia, entrando en operación en diciembre de 2015. Desde esa fecha, el sistema ha registrado la actividad solar de forma ininterrumpida, detectando cuatro potentes explosiones solares clase C presentadas los días 15, 16 y 17 del mes de abril de 2016. Las señales recibidas por el radio receptor son enviadas a la base de datos global del Stanford Solar Center, de Stanford University, con el código *UTP 0383*.

**Conclusiones:** El radio telescopio desarrollado permite al Observatorio Astronómico OAUTP adelantar el monitoreo del clima espacial y de la actividad solar, enviando la información recibida a repositorios de acceso público a nivel mundial. No existen otros receptores de radio de este tipo operando en la actualidad en países ecuatoriales, una región de la Tierra donde los efectos de la actividad solar pueden conducir a nuevos conocimientos que aporten en la comprensión de este fenómeno. Además, los filtros de orden superior diseñados pueden tener aplicación en el campo de la instrumentación para la medición de señales eléctricas biomédicas.

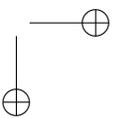
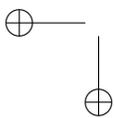
**Palabras clave:** Bancos de filtros, clima espacial, filtros Butterworth, fulguraciones solares, radiotelescopio.

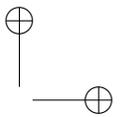
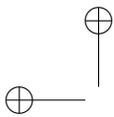
**Idioma:** Español.



Citación: A. F. Jaramillo, H. D. Galvis, E. A. Quintero, "Monitoreo del Clima Espacial desde Colombia Mediante Radio Receptores Butterworth de Orden Superior" INGENIERÍA, vol. 22, no. 1, pp. 83-97, 2017.

©Los autores; titular de derechos de reproducción Universidad Distrital Francisco José de Caldas. En línea DOI: <http://dx.doi.org/10.14483/udistrital.jour.reving.2017.1.a05>





### Abstract

**Context:** Solar flares and coronal mass ejections release great amounts of radiation to the outer space, altering the electrical properties of external layers in earth's atmosphere. The set of phenomena derived from this interaction is known as space weather, and its consequences include failures in the geolocation systems, telecommunications, satellite based operations, spatial tracking, radio navigation and overload of electrical network.

**Method:** This paper presents the development of a radio receptor enabling monitoring of space weather from Colombia. A novel methodology to design analog high-order Butterworth filters is described, based on the parallel interconnection of filtering banks of first and second order, yielding the realization of the desired transfer function. These filters are then used to build a radio receptor to monitor the mentioned solar activity.

**Results:** The radio receptor was installed in a space weather monitoring station in Colombia, specifically in the Astronomical Observatory of the Universidad Tecnológica de Pereira (OAUTP), becoming operational on December 2015. From that date, the radio receptor has been registering solar activity uninterruptedly; and it has been able to detect four powerful solar explosions of class C, registered on the days 15, 16 and 17th in the month of April 2016. The received signals are sent to the *Stanford Solar Center of Stanford University*, with the code *UTP 0383*.

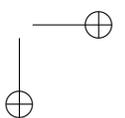
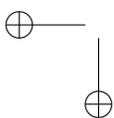
**Conclusions:** The newly developed radio telescope is enabling the Astronomical Observatory OAUTP to monitor space weather and solar activity, sensing and registering such information in global publicly available repositories. No other radio receptors of this kind are currently operating from equatorial countries, a region on Earth where the effects of solar activity may lead to new insights in understanding this phenomena. Besides, we anticipate that the designed high-order filters may have application as well in instruments for sensing biomedical electrical signals.

**Keywords:** Butterworth filters, filter banks, radio telescope, solar flares, space climate.

## 1. Introducción

La radioastronomía es la rama de la astronomía que estudia los cuerpos y fenómenos celestes a partir de la observación de la radiación emitida en la banda de radiofrecuencia. Dentro de este amplio espectro se encuentra la banda muy baja de frecuencias (VLF), situada entre 3 kHz y 30 kHz [1], en la cual emiten un amplio conjunto de estaciones situadas alrededor del mundo. Una de las radiodifusoras VLF más destacadas es NAA, también conocida como Cutler VLF Transmitter, debido a que posee un transmisor de 1 MW (el más potente en su tipo). Esta estación emite en 24 kHz y es utilizada por United States Navy para la comunicación de una sola vía con sus submarinos ubicados en el atlántico. Para cubrir las enormes distancias que se ven involucradas en la comunicación trasatlántica, esta estación, además de su gran potencia, hace uso de la ionósfera, de tal forma que la reflexión de la señal en esta capa atmosférica permite adaptar la comunicación a la curvatura de la Tierra [2].

Durante una fulguración solar se liberan al espacio enormes cantidades de partículas cargadas, las cuales al interactuar con la ionósfera perturban sus características eléctricas, generando interferencia en las señales VLF y, por ende, distorsión en la señal emitida por la estación NAA. Además, estas partículas afectan las comunicaciones satelitales, los sistemas de navegación aérea, el seguimiento de sondas espaciales, entre otros [3]. Al conjunto de efectos negativos que este tipo de

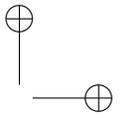
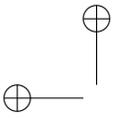


fenómenos solares causan en la actividad humana se le denomina clima espacial (Space Weather), el cual es monitoreado permanentemente por diferentes observatorios situados alrededor del mundo. Estos centros de investigación registran las señales emitidas por las estaciones VLF en busca de perturbaciones asociadas a fulguraciones solares, enviando sus registros a Stanford Solar Center de Stanford University, en donde son almacenadas en una base de datos global [20].

Con el fin de participar en esta red y en el marco de la línea de investigación en radioastronomía, en 2014 el Grupo de Investigación en Astroingeniería Alfa Orión con sede en el Observatorio Astronómico de la Universidad Tecnológica de Pereira (GIA-OAUTP) desarrolló un sistema para la captura, acondicionamiento, visualización y almacenamiento de la señal VLF emitida por la estación NAA, con el fin de monitorear el clima espacial desde Colombia, a partir de la identificación de las perturbaciones presentadas en esta señal [4]. El sistema construido se basó en el monitor de clima espacial SuperSID, desarrollado por Stanford University [5]. Sin embargo, este instrumento fue ideado para registrar cerca de siete estaciones VLF localizadas en Estados Unidos, con lo cual presenta un ancho de banda amplio (alrededor de 20 kHz). Dado que en nuestras latitudes solo es posible recibir la emisión en 24 kHz de la estación NAA (gracias a su gran potencia de transmisión), para el caso particular de Pereira, el ancho de banda nominal del instrumento construido se convierte en una fuente potencial de ruido.

En vista de la anterior problemática, en este artículo se presenta el diseño y la construcción de un radio receptor de banda estrecha centrado en 24 kHz que permita monitorear la señal emitida por la estación de muy baja frecuencia NAA, la cual es recibida por una antena tipo lazo rectangular de 2m de arista instalada en la terraza del Observatorio Astronómico de la Universidad Tecnológica de Pereira, eliminando de esta forma la interferencia observada en las señales arrojadas por el antiguo receptor, la cual se debe a las diversas señales electromagnéticas de diferentes frecuencias que son admitidas por el sistema debido a su amplio ancho de banda. El radio receptor desarrollado hace posible el monitoreo del clima espacial desde la posición geográfica ecuatorial de Colombia, arrojando registros que alimentan la base de datos global administrada por Stanford Solar Center de Stanford University, lo cual permite a la comunidad académica mundial dedicada al estudio de los fenómenos solares contar con datos de diferentes latitudes que amplíen el conocimiento existente actualmente acerca de cómo las partículas procedentes del sol interactúan con la ionósfera en diferentes latitudes y cómo se ven afectados los sistemas de comunicaciones y de transporte de energía eléctrica en todo el planeta.

Para el desarrollo del radio receptor de banda estrecha requerido en el OAUTP, centrado en la frecuencia particular de la estación NAA (24 kHz), es necesario recurrir a técnicas de diseño de filtros analógicos de orden superior, con respuestas máximamente planas, pero que permitan optimizar el proceso de construcción del esquema circuital resultante. Generalmente, la selectividad frecuencial de un sistema de filtrado se encuentra directamente ligada al orden de dicho sistema. Al aumentar el orden, la realización analógica se vuelve cada vez más engorrosa, por lo que en la mayoría de aplicaciones se recurre a una implementación digital. Sin embargo, las realizaciones digitales conllevan ciertas desventajas, entre las cuales se encuentran la pérdida inevitable de información debido al muestreo, la dificultad de procesar en tiempo real y la generación de ruido en altas frecuencias. Por esta razón, se decidió diseñar una metodología que permitiese implementar filtros analógicos de orden superior de una manera simple e intuitiva, mediante la paralelización de

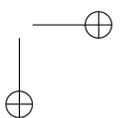
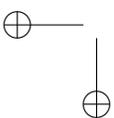


la función de transferencia y su implementación a partir de bancos de filtros basados en amplificadores operacionales.

La metodología diseñada hace uso de la visión sistémica, pues involucra la descomposición del filtro en subsistemas más simples contruidos a partir de elementos de uso común en electrónica analógica como transistores, arreglos analógicos programables, amplificadores de transconductancia y amplificadores operacionales. Por ejemplo, Rola A. Baki *et al.*, implementaron un filtro transistorizado de séptimo orden con un consumo no mayor a 13 mW y con una ganancia ajustable de hasta 13 dB [6]. Asimismo, en [7], Fabian Henrici *et al.*, desarrollaron una estrategia para el diseño de sistemas de filtrado a partir de arreglos analógicos programables (FPAA) y sus propias capacitancias parásitas internas. Por otra parte, Yao Li presenta en [8] el diseño de un filtro Chebyshev de cuarto orden basado en amplificadores de transconductancia, mientras que en [9] Randall L. Geiger y Edgar Sánchez presentan un tutorial para el diseño de filtros controlados por voltaje basados en este mismo tipo de amplificadores. Finalmente, en [10], Karima Garradhi *et al.*, proponen el diseño de un amplificador de transconductancia de bajo voltaje y bajo consumo de potencia para aplicaciones de filtros universales.

En cuanto a filtros basados en amplificadores operacionales se refiere, en [11] N. Neshatvar presenta el diseño de un filtro pasa bajos de tercer orden con frecuencia de corte de 100 Hz basado en redes Sallen-Key y pseudo-resistencias ajustables para aplicaciones en instrumentación biomédica. En [12], Saiteja Damera *et al.*, desarrollaron una técnica para el diseño de filtros mínimamente invasivos que introducen muy poco ruido a la señal a filtrar y que garanticen máxima linealidad, poniendo a prueba la estrategia a través del diseño de un filtro tipo Butterworth de quinto orden. Entre tanto, Ahmad Rezazadehrehyani *et al.*, plantean en [13] el diseño de un filtro Notch adaptativo con frecuencia de corte en 54.1 MHz, basado en el principio de cancelación de ruido. Asimismo, Brian Drost *et al.*, proponen en [14] el diseño de un filtro pasa bajos de cuarto orden mediante una estrategia basada en integradores configurados como osciladores circulares, destacándose que el prototipo final ocupa un área de tan solo 0.29 mm<sup>2</sup>. Finalmente, Chi-Hsiang lo presenta en [15] el diseño de un filtro pasa bajos de 4<sup>o</sup> orden con frecuencia de corte de 750 MHz, para aplicaciones anti-aliasing, y a partir de redes Sallen-Key.

Sin embargo, los sistemas de filtrado registrados en el estado del arte presentan la desventaja de que están basados en metodologías aplicables únicamente a frecuencias muy específicas, alejándolos de aplicaciones particulares con frecuencias poco comunes, como es el caso de la estación de baja frecuencia NAA. Además, los elementos involucrados en estos diseños y su posterior ensamble exigen tecnologías de elaboración de circuitos impresos de difícil acceso y que, por lo tanto, demandan altos costos de desarrollo. De esta forma, con el fin de construir el radio receptor de banda estrecha centrado en 24 kHz presentado en este artículo, se desarrolló una novedosa estrategia alternativa para el diseño y la realización circuital de filtros activos analógicos de orden superior, basada en la descomposición de la función de transferencia original en subsistemas de primero y segundo orden, implementados a partir de integradores y redes Sallen-Key basados en amplificadores operacionales de uso común y bajo costo.



## 2. Materiales y métodos

Para llevar a cabo la implementación de un sistema de filtrado es necesario partir de su función de transferencia, la cual tiene la forma:

$$H(s) = \frac{A(s)}{B(s)} \quad (1)$$

Donde la información de respuesta transitoria está en  $A(s)$ , mientras que la información de respuesta estacionaria se encuentra en  $B(s)$ . De allí entonces:

$$H(s) = \frac{A_0(s)}{B_0(s)} + \frac{A_1(s)}{B_1(s)} + \dots + \frac{A_n(s)}{B_n(s)} = \frac{A(s)}{B(s)} \quad (2)$$

$$H(s) = H_0(s) + H_1(s) + \dots + H_n(s) \quad (3)$$

$$H(s) = \sum_{i=0}^n H_i(s) \quad (4)$$

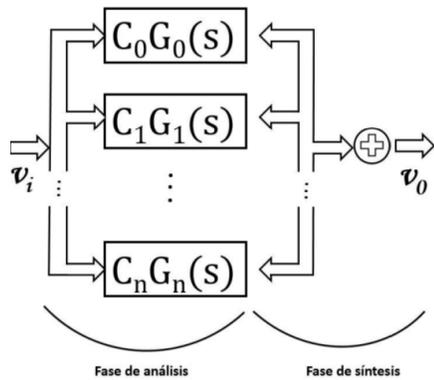
Generalizando, el resultado anterior se escribe como:

$$H(s) = \sum_{i=0}^n G_i(s) \quad (5)$$

con:

$$G_i(s) = C_i H_i(s) \quad (6)$$

Cada  $G_i(s)$  posee infinitas realizaciones, lo cual ofrece gran versatilidad a la hora del diseño y permitiendo así que la reconstrucción de  $H(s)$  se pueda realizar mediante bancos de filtros. En general,  $H(s)$  representa al sistema en el espacio  $s$  y está dada por la sumatoria (5), donde cada  $G_i(s)$  se considera como un subsistema, y a  $H(s)$  como una superposición de las respuestas en frecuencia de estos subsistemas individuales. Una consideración particular es el caso de los sistemas lineales e invariantes en el tiempo, pues dado que todos los coeficientes de los polinomios encontrados en  $H(s)$  son reales, y que dicho sistema es propio, entonces es claro que cada  $G_i(s)$ , al descomponer  $H(s)$  hasta donde sea posible, debe ser un subsistema de primero o segundo orden con polos complejos conjugados.



**Figura 1.** Esquema general de la metodología para la reconstrucción de sistemas mediante bancos de filtros

En vista de lo anterior, la metodología propuesta plantea los siguientes pasos como parte del algoritmo de reconstrucción de  $H(s)$  mediante bancos de filtros  $G_i(s)$ :

1. Obtener la función de transferencia  $H(s)$  del filtro a implementar.
2. Hallar cada  $H_i(s)$ .
3. Determinar los coeficientes de flexibilidad  $C_i$  que actúan como coeficientes de los subsistemas  $H_i(s)$ , y los subsistemas  $G_i(s)$ , para de esta manera implementarlos de forma independiente como sistemas de primero o segundo orden.
4. Reconstruir el sistema dado por  $H(s)$  mediante la combinación lineal de los diferentes  $G_i(s)$  a través de sumadores analógicos con amplificadores operacionales.

Estos pasos se resumen en forma de diagrama de bloques en la Figura 1.

Al tratarse de sistemas de primero y segundo orden, cada  $G_i(s)$  puede implementarse mediante integradores, redes *Sallen-Key* o realimentadores de estados [16], cuyos diagramas esquemáticos se presentan en las figuras 2, 3 y 4 respectivamente.

La metodología desarrollada se puso a prueba mediante el diseño de un filtro de orden superior para el sistema de monitoreo del clima espacial del Observatorio Astronómico de la Universidad Tecnológica de Pereira. Los requerimientos de diseño involucran un filtro pasa banda con frecuencia central de 24 kHz (frecuencia de emisión de la estación NAA) y con una ganancia en banda pasante de 20 dB. Lo anterior con el fin de que, de la totalidad de las señales detectadas por la antena lazo rectangular, solo la emitida por la estación NAA sea efectivamente registrada por el software de procesamiento con una amplitud aceptable. En vista de lo anterior, utilizando Matlab se obtuvo la función de trans-

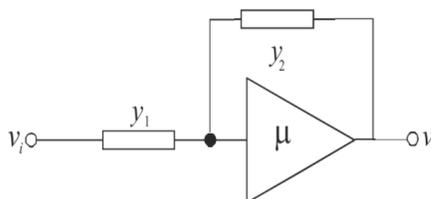


Figura 2. Realización de un sistema de primer orden a partir de un amplificador operacional.

ferencia del filtro requerido (ecuación 7), la cual es de características Butterworth de sexto orden, con ganancia en banda pasante de 20 dB y frecuencias de corte lateral superior y lateral inferior de 26 KHz y 24 KHz respectivamente.

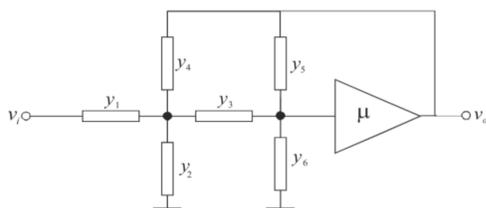


Figura 3. Realización de un sistema de segundo orden a través de redes *Sallen-Key*.

$$H(s) = \frac{1,984 \cdot 10^{13} s^3}{s^6 + 2,513 \cdot 10^4 \cdot s^5 + 7,422 \cdot 10^{10} \cdot s^4 + 1,24 \cdot 10^{15} \cdot s^3 + 1,828 \cdot 10^{21} \cdot s^2 + 1,525 \cdot 10^{25} \cdot s + 1,495 \cdot 10^{31}} \quad (7)$$

Al aplicar la estrategia de realización descrita en párrafos anteriores, se obtuvieron tres subsistemas,  $H_1(s)$ ,  $H_2(s)$  y  $H_3(s)$ , todos ellos de segundo orden, los cuales se presentan en las ecuaciones (8), (9) y (10) respectivamente.

$$H_1(s) = \frac{-31087s + 5,273 \cdot 10^8}{s^2 + 8050 \cdot s + 2,727 \cdot 10^{10}} \quad (8)$$

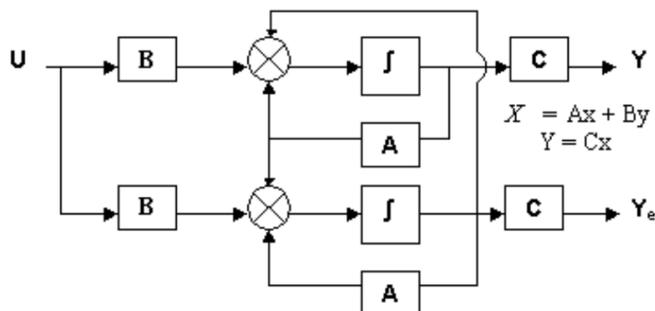


Figura 4. Implementación de un sistema de segundo orden a partir de un realimentador de estados.

$$H_2(s) = \frac{21974s + 5,2310^9}{s^2 + 15000 \cdot s + 2,355 \cdot 10^{10}} \quad (9)$$

$$H_3(s) = \frac{9113s - 5,621 \cdot 10^9}{s^2 - 1888 \cdot s + 2,328 \cdot 10^{10}} \quad (10)$$

Para simplificar la realización circuital de estos subsistemas se optó por eliminar el componente de  $s^0$  de cada numerador, pues esta situación no generó cambios significativos en el diagrama de Bode de la función de transferencia original. Las funciones de transferencia de los sistemas finalmente implementadas se presentan en las ecuaciones (11), (12) y (13) donde se realizó además una normalización para compensar la ganancia de la red Sallen- Key. Estos filtros deben interconectarse como se ilustra en la Figura 1, dando lugar al banco de filtros. En la ecuación (14) se muestra el sistema final implementado, el cual contiene los cambios mencionados sobre las funciones de transferencia originales de los subsistemas.

$$H_1(s) = \frac{-93261s}{s^2 + 8050 \cdot s + 2,727 \cdot 10^{10}} \quad (11)$$

$$H_2(s) = \frac{65922s}{s^2 + 15000 \cdot s + 2,355 \cdot 10^{10}} \quad (12)$$

$$H_3(s) = \frac{27339s}{s^2 - 1888 \cdot s + 2,328 \cdot 10^{10}} \quad (13)$$

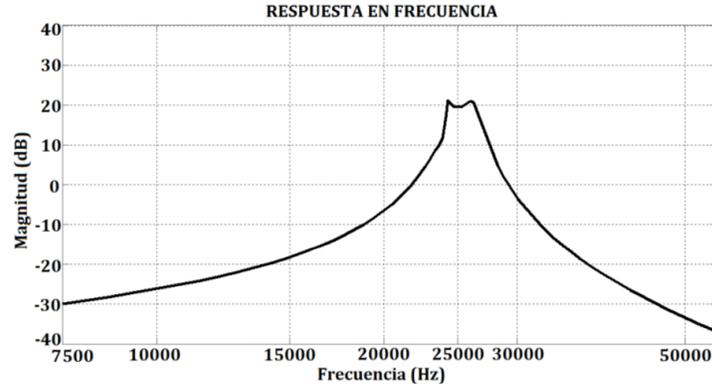
$$H(s) = \frac{1,864 \cdot 10^8 s^4 + 3,588 \cdot 10^{14} s^3 - 3,091 \cdot 10^{18} s^2 + 8,268 \cdot 10^{24} s}{s^6 + 2,116 \cdot 10^4 \cdot s^5 + 7,417 \cdot 10^{10} \cdot s^4 + 1,039 \cdot 10^{15} \cdot s^3 + 1,827 \cdot 10^{21} \cdot s^2 + 1,272 \cdot 10^{25} \cdot s + 1,495 \cdot 10^{31}} \quad (14)$$

De lo anterior y de acuerdo a lo planteado en la ecuación (6), se concluye que:

$$C_1 = C_2 = C_3 = 1 \quad (15)$$

De esta forma, el filtro  $H(s)$  de la ecuación (14) se obtiene calculando:

$$H(s) = C_1 H_1(s) + C_2 H_2(s) + C_3 H_3(s) \quad (16)$$



**Figura 5.** Diagrama de *Bode* del filtro *Butterworth* pasa banda de sexto orden diseñado para el sistema de monitoreo del clima espacial en Colombia con sede en el Observatorio de la Universidad Tecnológica de Pereira

El diagrama de *Bode* del filtro de la ecuación (14) se observa en la Figura 5. Nótese que se trata de un filtro pasa banda de alta selectividad, centrado alrededor de 25 kHz.

Al tratarse de sistemas de segundo orden, su implementación se realizó a través de redes *Sallen-Key* (Figura 3), las cuales se describen a través de la matriz definida de admitancias presentada en la ecuación (17), de acuerdo a la teoría de la descripción matricial de redes electrónicas activas [16].

$$MDA = \begin{bmatrix} y_1 & -y_1 & 0 & 0 \\ -y_1 & y_1 + y_2 + y_3 + y_4 & -y_3 & -y_4 \\ 0 & -y_3 & y_1 + y_5 + y_6 & -y_5 \\ 0 & -y_4 & -y_5 & y_4 + y_5 \end{bmatrix} \quad (17)$$

De la anterior matriz se obtiene que la función de transferencia de la red *Sallen-Key* está dada por [16]:

$$H(s) = \frac{\mu y_1 y_3}{(y_1 + y_2 + y_4)(y_3 + y_5 + y_6) + y_3(y_5 + y_6) - \mu y_5(y_1 + y_2 + y_3 + y_4) + y_3 y_4} \quad (18)$$

Al igualar la función de transferencia general de la red *Sallen-Key* presentada en la ecuación (18) con las funciones de transferencia deseadas dadas por las ecuaciones (11), (12) y (13), se obtuvieron los circuitos que se presentan en las Figuras 6, 7 y 8, respectivamente. Finalmente, la Figura 9 presenta el sumador final que permite realizar el banco de filtros de acuerdo a lo indicado en la Figura 1, mientras que la Figura 10 ilustra la respuesta en frecuencia del sistema obtenida en *Proteus*.

### 3. Resultados

El sistema desarrollado fue instalado el mes de diciembre del año 2015 en el Observatorio Astronómico de la Universidad Tecnológica de Pereira, en reemplazo del antiguo radio receptor utilizado para adelantar el monitoreo del clima espacial desde Colombia. Para la instalación, la entrada

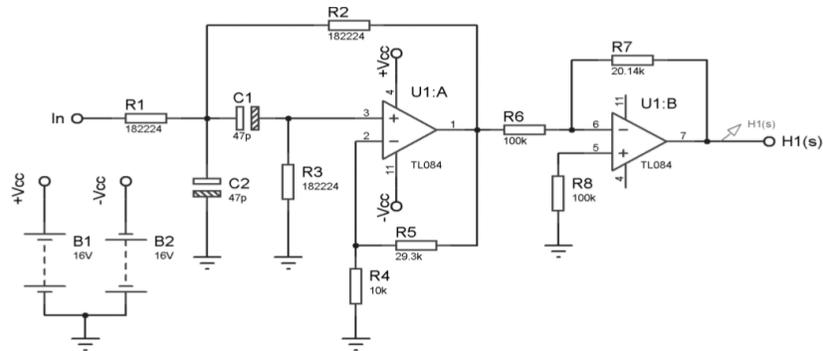


Figura 6. Implementación de la función de transferencia  $H_1(s)$  dada por la ecuación (11).

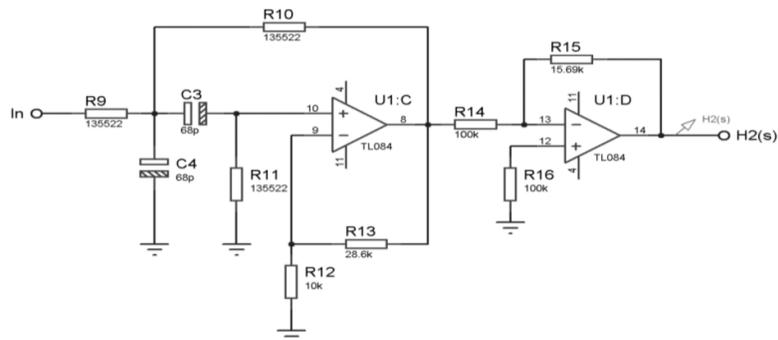


Figura 7. Implementación de la función de transferencia  $H_2(s)$  dada por la ecuación (12).

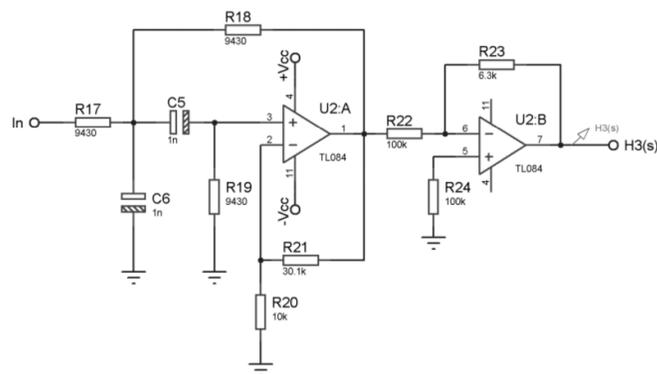


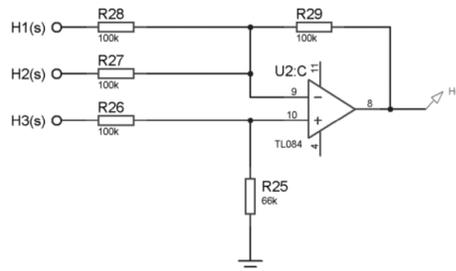
Figura 8. Implementación de la función de transferencia  $H_3(s)$  dada por la ecuación (13).

del filtro construido se conectó a la antena tipo lazo rectangular de 2m de arista existente en la terraza del Observatorio, mientras que la salida se incrustó en la entrada auxiliar de la tarjeta de sonido de un computador personal (con frecuencia de muestro de 96 kbps), en el cual la aplicación

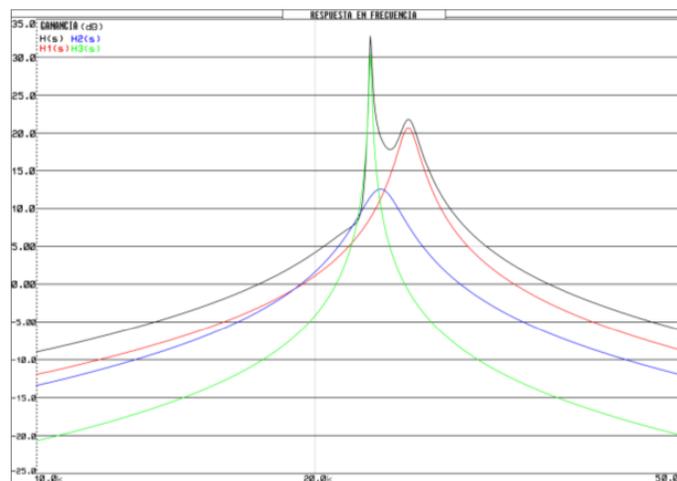
*MonitorCEC* (desarrollada por personal del Observatorio Astronómico) captura, visualiza y almacena la actividad detectada en una base de datos local [22]. Este software también se encarga de enviar diariamente la señal recibida al Stanford Solar Center de Stanford University, con el fin de alimentar la base de datos global gestionada por este organismo.

Desde la entrada en operación del filtro desarrollado se han registrado seis explosiones solares clase C, de acuerdo a la escala de intensidad establecida por National Oceanic and Atmospheric Administration (NOAA, por sus siglas en inglés). Las Figuras 11, 12 y 13 presentan las tres detecciones más potentes registradas por el sistema desde su entrada en operación, tomadas directamente de la base de datos de Stanford Solar Center de Stanford University [23].

La Figura 14 permite observar el aspecto final del radio receptor desarrollado, la antena tipo lazo rectangular con la que se cuenta en el Observatorio Astronómico de la UTP, y el programa SuperSID de Stanford Solar Center, que adelanta la captura, visualización y almacenamiento de los datos recibidos. Entre tanto, la Figura 15 ilustra la densidad espectral de potencia de la señal capturada por el antiguo radio receptor (arriba) y del presentado en este trabajo (abajo). Las elipses rojas identifican el armónico en 24 kHz correspondiente a la señal de baja frecuencia emitida por la estación NAA. Nótese que en el nuevo radio receptor este armónico es el más intenso de todo el espectro, alcanzando una potencia de 56.321 dB, en contraste con los 32.517 dB arrojados por el antiguo sistema basado en el esquema de Stanford Solar Center.



**Figura 9.** Sumador que permite obtener de la función de transferencia  $H(s)$ .



**Figura 10.** Respuesta en frecuencia del filtro pasa banda Butterworth de sexto orden diseñado. Negro:  $H(s)$ . Rojo:  $H_1(s)$ . Azul:  $H_2(s)$ . Verde:  $H_3(s)$ .

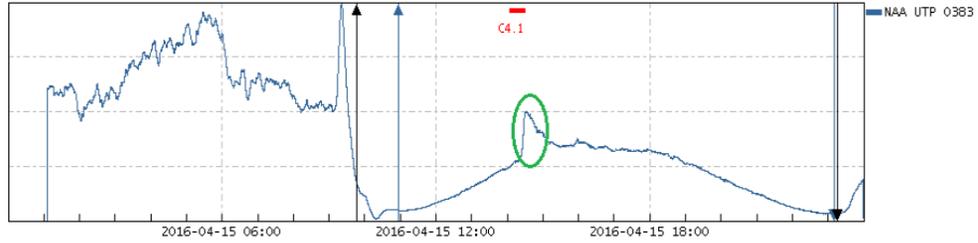


Figura 11. Explosión solar clase C4.1 presentada el 15 de abril de 2016 a las 14:04UT.

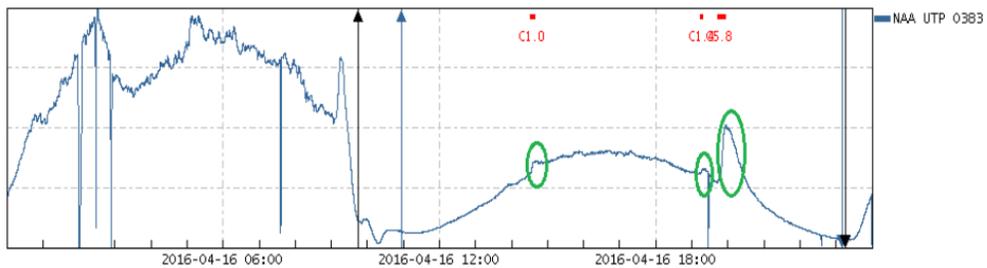


Figura 12. Explosiones solares clase C1.0, C1.4 y C5.8 presentadas el 16 de abril de 2016.

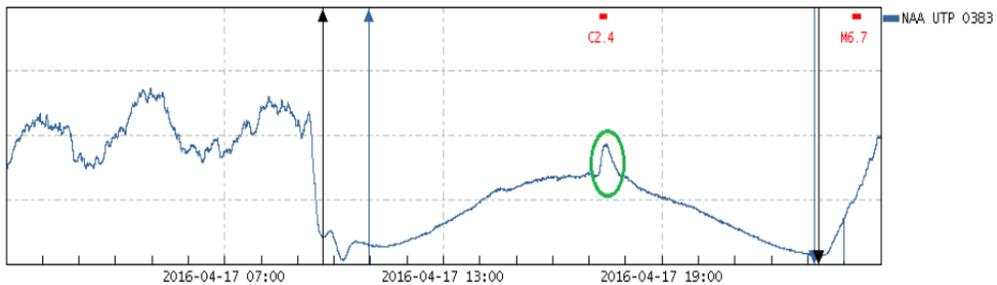
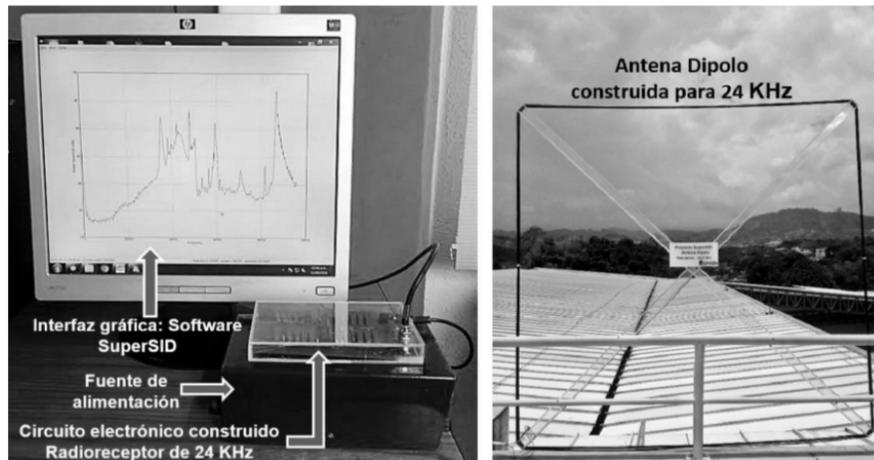


Figura 13. Explosión solar clase C2.4 presentada el 17 de abril de 2016.

## 4. Conclusiones

La metodología desarrollada para la implementación de filtros activos analógicos de orden superior a través de la descomposición de su función de transferencia, permite realizar sistemas mediante la interconexión en paralelo de montajes circuitales elementales de primero y segundo orden, lo cual reduce significativamente la complejidad de la realización, además de posibilitar una



**Figura 14.** Izquierda: Aspecto del software *SuperSID* y del radioreceptor desarrollado. Derecha: Fotografía de la antena tipo lazo rectangular de 2 m de arista mediante la cual se realiza la captura de la señal emitida por la estación *NAA* en 24 kHz.

reconstrucción mucho más fiel a la respuesta en frecuencia deseada, al prescindirse de la conexión sucesiva de subsistemas en cascada. Esta situación abre las puertas para que la estrategia presentada en este artículo sea implementada en un sinnúmero de aplicaciones en las que se requiere adelantar un filtrado analógico de alta selectividad. Entre otros campos de aplicación, es posible citar el desarrollo de sistemas de filtrado para la eliminación de armónicos indeseados en las redes de energía eléctrica [17], la medición de variables físicas en donde se requiere minimizar el ruido aditivo involucrado en el proceso de transducción de la señal, o en la instrumentación biomédica, área en la que permanentemente se requieren filtros de ranura de alta velocidad de caída que permitan aislar la señal de interés del componente indeseado de 50/60 Hz. Por ejemplo, en este último campo de aplicación, la estrategia desarrollada podría ser utilizada para el diseño de filtros activos analógicos de alta selectividad que permitan eliminar la componente de 60 Hz presente en las señales electrocardiográficas registradas por sistemas electrónicos que se alimentan de la red eléctrica.

Por otra parte, el radio receptor construido permitió actualizar el sistema para el monitoreo del clima espacial desde Colombia existente en el Observatorio Astronómico de la Universidad Tecnológica de Pereira, al involucrar un sistema de filtrado adecuado a la señal de la estación *NAA* emitida en 24 kHz, siendo esta la estación VLF recibida con mayor intensidad en la posición geográfica en la que se encuentra el país. De esta forma, el sistema desarrollado se adecúa a las condiciones particulares presentes en el Observatorio, arrojando registros con una relación señal a ruido muy superior, pues el antiguo sistema, desarrollado por Stanford University, está diseñado para operar en latitudes septentrionales, por lo cual posee un ancho de banda amplio que permite registrar señales de al menos 20 estaciones VLF, pero que, al ser instalado en el Observatorio de la UTP, presentó una alta vulnerabilidad al ruido.

Asimismo, es necesario mencionar que el presente trabajo fortaleció significativamente la línea de investigación en radioastronomía, haciendo posible el desarrollo de proyectos subsecuentes que

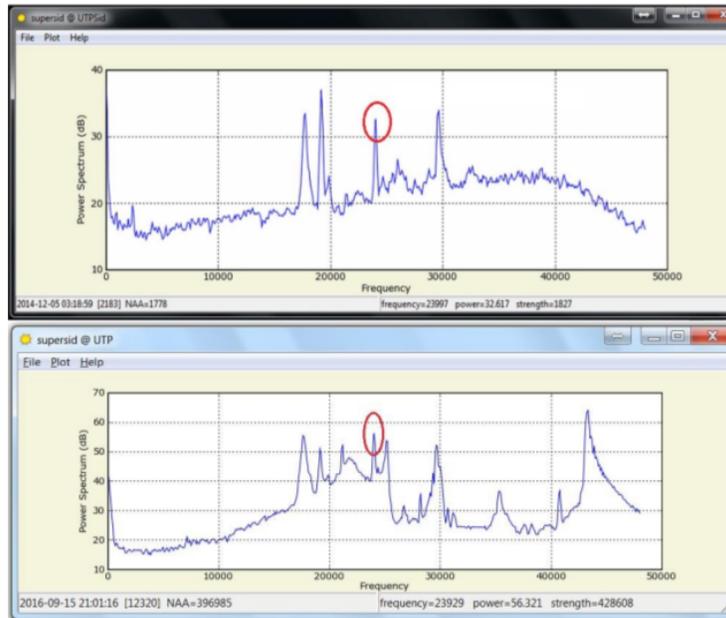


Figura 15. Densidad espectral de potencia del antiguo radio receptor (arriba) y del presentado en este artículo (abajo).

buscan registrar los fenómenos solares de forma directa en otras bandas de frecuencia, como por ejemplo 20,1 MHz [18]. Además, el sistema desarrollado proporciona, a través de la base de datos de Stanford Solar Center, información acerca de las perturbaciones producidas por la actividad solar en la ionosfera terrestre, pero detectadas desde latitudes ecuatoriales como en las que se ubica Colombia, complementando de esta forma la red global de observaciones del clima espacial, gracias a la cual investigadores de todo el mundo tienen la posibilidad de ampliar el conocimiento que se tiene acerca de este fenómeno. Por ejemplo, vale la pena mencionar que en [19] se presenta la correlación existente entre las fulguraciones solares y las perturbaciones súbitas en la ionosfera terrestre (SID) registradas al medir la variabilidad presentada en la intensidad de las ondas de radio de muy baja frecuencia (VLF) reflejadas en la capa D de la ionosfera, para lo cual los autores hacen uso de datos del monitor del clima espacial presentado en este artículo y del existente en el Observatorio Astronómico Centroamericano de Suyapa de la Universidad Nacional Autónoma de Honduras.

Finalmente, como trabajo futuro es necesario recalcar que, si bien la metodología involucra la descomposición de la función de transferencia del filtro deseado en subsistemas de primero y segundo orden, y que el filtro propuesto como ejemplo cabe dentro de este esquema, no todas las funciones de transferencia pueden descomponerse de esta forma, pues en aquellas que presenten la forma de la ecuación (19) es inevitable la aparición de sistemas de tercer orden, además de los de primero y segundo.

$$H(s) = \frac{A(s)}{(s \pm a)^n * B(s)} \text{ con } n > 2 \quad (19)$$

En vista de esta situación, es necesario involucrar esta variante del método en la formulación del mismo, además de presentar el esquema general del circuito de un sistema de tercer orden construido a partir de realimentadores de estados, de tal forma que la implementación de la estrategia desarrollada sea independiente de la aplicación donde se desee usar y de la función de transferencia del filtro requerido.

### Agradecimientos

Los autores manifiestan expresamente sus agradecimientos a la Vicerrectoría de Investigaciones, Innovación y Extensión de la Universidad Tecnológica de Pereira por la financiación del proyecto “Sistema para el Monitoreo del Clima Espacial en Colombia”, código 3-14-1, de cuyos resultados se desprende el presente artículo.

### Referencias

- [1] S. Manoj y S. Sundaram, “Dynamic Spectrum of Long Wave Terrestrial Radio Signals During Episodes of Ionospheric Disturbances Caused By Solar Activity”. *Journal of Chemical and Pharmaceutical Sciences*, Volumen 9, Número 1, 2016, pp. 548-553. ↑84
- [2] P. Kulkarni, U. S. Inan, T. F. Bell y J. Bortnik, “Precipitation Signatures of Ground-Based VLF Transmitters”. *Journal of Geophysical Research: Space Physics*, Volumen 113, Número A7, 2008, pp. 1-11. ↑84
- [3] A. Petry, J. R. de Souza, H. F. de Campos Velho, A. G. Pereira y G. J. Bailey, “First Results of Operational Ionospheric Dynamics Prediction for the Brazilian Space Weather Program”. *Advances in Space Research*, Volumen 54, Número, 2014, pp. 22-36. ↑84
- [4] L. M. Puerta, A. C. Acuña, O. E. Pulgarín y E. A. Quitero. “Sistema para el Registro de Llamadas Solares en la Banda Muy Baja de frecuencias”. *Revista de Investigaciones Universidad del Quindío*, Volumen 1, Número 26, 2014, pp. 103-111. ↑85
- [5] D. Scherrer, R. Mitchell, T. Huynh, W. Lord y M. Lord, *SuperSID Manual – Space Weather Monitors*, 2009. [En línea]. Disponible en: <http://solarcenter.stanford.edu/SID/Distribution/SuperSID/supersid.v1.1/Doc/SuperSIDManual.v1.pdf> ↑85
- [6] R. A. Baki y M. N. El-Gamal, “A Low-Power, 5-70MHz, 7th-Order Filter with Programmable Boost, Group Delay, and Gain Using Instantaneous Companding”. *Low Power Electronics and Design International Symposium*, 2001, pp. 299-304. ↑86
- [7] F. Henrici, J. Becker, A. Buhmann, M. Ortmanns y Y. Manoli, “A Continuous-Time Field Programmable Analog Array Using Parasitic Capacitance Gm-C Filters”. *Chair of Microelectronics, Circuits and System ISCAS IEEE International Symposium*, 2007, pp. 2236-2239. ↑86
- [8] Y. Li, X. Zhang, L. Zhu y W. Song, “The Study of Analog Evolvable Chebyshev Transconductance Filters”. *World Congress on Intelligent Control and Automation*, 2008, pp. 4531-4536 ↑86
- [9] R. L. Geiger y E. Sánchez, “Active Filter Design Using Operational Transconductance Amplifiers: A Tutorial”. *IEEE Circuits and Devices Magazine*, pp 20- 32, 1985. ↑86
- [10] K. Garradhi, N. Hassen y K. Besbes, “Low Voltage Low Power Analog Circuit Design OTA Using Signal Attenuation Technique in Universal Filter Application”. *12th International Multi-Conference on Systems, Signals & Devices*, 2015, pp. 1-7. ↑86
- [11] N. Neshatvar, “Design of Low Frequency Analog Low Pass Filter Using Tunable Pseudo Resistors”. *Middle East Conference on Biomedical Engineering (MECBME)*, 2014, pp. 39-42. ↑86
- [12] S. Damera, A. Karsilayan y J. Silva, “Design of Minimally-Invasive All-Pole Analog Lowpass Filters”. *IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2014, pp. 1025-1028. ↑86
- [13] A. RezazadehReyhani, C. Jayanthmurthy, B. Gillman, J. Walling, J. Belz y B. Farhang, “An Analog Adaptive Notch Filter Based on the Noise Cancellation Principle”. *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2015, pp. 2660-2663. ↑86

- [14] B. Drost, M. Talegaonkar y P. K. Hanumolu, "Analog Filter Design Using Ring Oscillator Integrators". *IEEE Journal of Solid-State Circuits*, Volumen 47, Número 12, 2012, pp. 3120-3129. ↑86
- [15] Chi-Hsiang Lo, "A 6 mW 750 MHz anti-aliasing filter for receiver analog front-end". *Analog Integrated Circuits and Signal Processing*, Volumen 79, Número 3, 2014, pp. 599-604. ↑86
- [16] L. E. Avendaño. *Sistemas Electrónicos Analógicos: Un Enfoque Matricial*. Editorial Universidad Tecnológica de Pereira, Colombia, 2007, pp. 420-423. ↑88, 90
- [17] H. E. Rojas, E. Rivas y A. A. Jaramillo. "Aspectos técnicos y normativos para el monitoreo y medición de armónicos", *Ingeniería*, Volumen 19, Número 2, 2014, pp. 129-146. ↑94
- [18] H. D. Galvis, E. A. Quintero y L. F. Cardona, "Development of a Magnetic Loop Antenna for the Detection of Jovian Radiowaves at 20.1 MHz". *Tecciencia*, Volumen 11, Número 20, 2016, pp. 21-46. ↑95
- [19] A. Artola, E. Mills, y P. Estrada, "Correlación entre la variabilidad en la banda VLF en Francisco Morazán y la actividad solar," *Ciencias Espaciales*, Volumen 6, Número 1, 2016, pp. 1-24. ↑
- [20] Space Weather Monitors, *Stanford Solar Center*. S.f. [En línea]. Disponible en: <http://sid.stanford.edu/database-brow-ser/browse.jsp?date=2016-04-06T00.00.00> ↑85
- [21] Space Weather Monitors, *Stanford Solar Center*, UTP 0383, S.f. [En línea]. Disponible en: <http://sid.stanford.edu/database-brow-ser/browse.jsp> ↑
- [22] Monitor CEC, S.f. [En línea]. Disponible en: <https://drive.google.com/open?id=0By4tA9RUXwKuaFNiCkFkcDFoSIU> ↑92
- [23] Stanford Solar Center de Stanford University, S.f. [En línea]. Disponible en: <https://drive.google.com/open?id=0By4tA9RUXwKuaFNiCkFkcDFoSIU> ↑92

### Andrés Felipe Jaramillo Alvarado

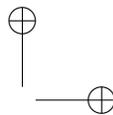
Nació en Pereira, Colombia. Es Ingeniero Físico de la Universidad Tecnológica de Pereira, Colombia. Actualmente cursa estudios de Maestría en Ciencias Computacionales en el Instituto Nacional de Astronomía, Óptica y Electrónica de México (INAOE), Puebla, México. Se desempeñó durante 3 años como estudiante investigador en el Grupo de Investigación en Astroingeniería Alfa Orión con sede en el Observatorio Astronómico de la Universidad Tecnológica de Pereira, participando en el proyecto "Sistema para el Monitoreo del Clima Espacial en Colombia", financiado por la Vicerrectoría de Investigaciones, Innovación y Extensión de la Universidad Tecnológica de Pereira.  
e-mail: anfejaramillo@utp.edu.co

### Hamilton David Galvis Rodríguez

Nació en Medellín, Colombia. Es Ingeniero Físico de la Universidad Tecnológica de Pereira, Colombia. Actualmente cursa estudios de Maestría en Instrumentación Física en la Universidad Tecnológica de Pereira, Colombia. Se desempeñó durante 3 años como estudiante investigador en el Grupo de Investigación en Astroingeniería Alfa Orión con sede en el Observatorio Astronómico de la Universidad Tecnológica de Pereira, participando en el proyecto "Sistema para el Monitoreo del Clima Espacial en Colombia", financiado por la Vicerrectoría de Investigaciones, Innovación y Extensión de la Universidad Tecnológica de Pereira.  
e-mail: hdbgavis@utp.edu.co

### Edwin Andrés Quintero Salazar

Nació en Bogotá, D.C., Colombia. Es Ingeniero Electrónico de la Universidad Nacional de Colombia. Obtuvo su título de Especialización en Pedagogía para la Docencia Universitaria en la Fundación Universitaria del Área Andina, y de Maestría en Instrumentación Física en la Universidad Tecnológica de Pereira, Colombia. Actualmente cursa estudios de Maestría en Astronomía y Astrofísica en la Universidad Internacional de Valencia, España. Actualmente se desempeña como profesor en el área de instrumentación física en la Facultad de Ingenierías de la Universidad Tecnológica de Pereira, Colombia, y es director del Grupo de Investigación en Astroingeniería Alfa Orión con sede en el Observatorio Astronómico de la misma Universidad. Así mismo, fue investigador principal del proyecto "Sistema para el Monitoreo del Clima Espacial en Colombia", financiado por la Vicerrectoría de Investigaciones, Innovación y Extensión de la Universidad Tecnológica de Pereira, de cuyos resultados parciales se desprende el presente artículo.  
e-mail: equintero@utp.edu.co



## Investigación

**Evaluación de los Parámetros Físicoquímicos en un Reactor Discontinuo de Lodos Activados para el Tratamiento de Aguas con Metanol***Evaluation of Physicochemical Parameters in Discontinuous Reactors of Activated Sludge for the Treatment of Waters with Methanol***Marvin Caravali<sup>1</sup>, Martha L. Lote<sup>1</sup>, Lena Carolina Echeverry\*<sup>1</sup>**<sup>1</sup>Programa de Ingeniería Ambiental. Universidad Distrital Francisco José de Caldas. Bogotá - Colombia

\*Correspondencia: lcecheverry@udistrital.edu.co

Recibido: 05-07-2016. Modificado: 18-11-2016. Aceptado: 15-12-2016

## Resumen

**Contexto:** En la actualidad existen varios tipos de tecnologías limpias que ayudan a la descontaminación del agua residual industrial, como los sistemas de lodos activados. Dichas tecnologías han resultado efectivas en la remoción de materia orgánica y coloidal en aguas residuales domésticas e industriales, biosorción de metales pesados y remoción de patógenos y nutrientes (N y P). Lo anterior motiva la propuesta de trabajar con un sistema de tratamiento de lodos activados a escala en un ambiente controlado, a fin de demostrar su efectividad en tratamiento de aguas con metanol.

**Método:** El sistema de tratamiento de lodos activados se diseña y construye a escala a nivel de laboratorio para estudiar diferentes concentraciones de metanol en aguas residuales de una industria. En los montajes de cinco sistemas de tratamiento se realizan mediciones de parámetros físicoquímicos como: pH, oxígeno disuelto, temperatura, demanda química de oxígeno, sólidos suspendidos totales y la pérdida de metanol; a las 0, 24 y 48 h por cuadruplicado para observar el comportamiento de dichos parámetros durante la reducción de metanol.

**Resultados:** Para los sistemas de lodos activados se demuestra una pérdida promedio de metanol del 10,3 %. Los promedios de oxígeno disuelto decrecieron en el tiempo, aunque la demanda química de oxígeno es variable. Cada sistema con concentración diferente de metanol presenta comportamiento particular.

**Conclusiones:** Esta investigación indica que la biodegradación de metanol en aguas residuales industriales es posible, teniendo en cuenta que en el sistema de tratamiento a escala se obtuvo una disminución de metanol en el tiempo. Además, los resultados motivan a continuar investigando sobre esta tecnología limpia, de la cual no se dispone suficiente literatura en lo que respecta al manejo de lodos activados para la eliminación de metanol.

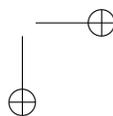
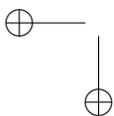
**Palabras clave:** Metanol físicoquímica, lodos activados, metanol, reactor discontinuo, tratamiento de aguas.

**Agradecimientos:** Los autores agradecen al Centro de Investigaciones y Desarrollo científico (CIDC) de la Universidad Distrital Francisco José de Caldas por la financiación del proyecto código: 3-180-328-12. Igualmente agradecen al Sena-TecnoParque nodo Bogotá D.C., por permitir el desarrollo de este proyecto en sus instalaciones.

**Idioma:** Español

Citación: M. Caravali, M. L. Lote, L. C. Echeverry, "Evaluación de los Parámetros Físicoquímicos en un Reactor Discontinuo de Lodos Activados para el Tratamiento de Aguas con Metanol" INGENIERÍA, vol. 22, no. 1, pp. 98-110, 2017.

©Los autores; titular de derechos de reproducción Universidad Distrital Francisco José de Caldas.

En línea DOI: <http://dx.doi.org/10.14483/udistrital.jour.reveng.2017.1.a06>

### Abstract

**Context:** There are currently several types of clean technologies that help the decontamination of industrial wastewater, such as activated sludge systems. These technologies have proved effective in the removal of organic matter and colloidal in domestic wastewater and industrial, biosorption of heavy metals, and removal of pathogens and nutrients (N and P). It previous motivates the proposal of work with a system of treatment of sludge activated to scale in an environment controlled, to demonstrate its effectiveness in treatment of waters with methanol.

**Method:** The treatment system of activated sludge is designed and built to scale in laboratory level to study different concentrations of methanol in waste water from one industry. In assembly of five treatment systems is measured physicochemical parameters as: pH, dissolved oxygen, temperature, chemical oxygen demand, total suspended solids, and loss of methanol, at 0, 24 and 48 h by quadruplicate to observe the behavior of these parameters during the reduction of methanol.

**Results:** For activated sludge systems is demonstrated an average loss of 10.3 % methanol. The averages of dissolved oxygen decrease in the time although chemical oxygen demand is variable. Each system with different concentration of methanol presents behavior.

**Conclusions:** This research indicates that the biodegradation of methanol in industrial wastewater is possible, bearing in mind that the treatment system to scale was obtained a decrease of methanol in the time. In addition, the results motivate continue researching this clean technology, which does not have sufficient literature about the handling of activated sludge to the elimination of methanol.

**Keywords:** Activated sludge, batch reactor, methanol, physicochemical, water treatment

**Acknowledgements:** The authors give thanks to Research and Scientific Development Center (CIDC) of Universidad Francisco José de Caldas for financing this Project with Code: 3-180-328-12. Moreover, the authors give thanks to SENA-TecnoParque Node Bogota DC for allowing the development of this Project in its facilities.

## 1. Introducción

El metanol ( $CH_3OH$ ) es un alcohol utilizado en la fabricación de una amplia gama de derivados químicos como: metilaminas, formaldehído, ácido acético, metacrilato de metilo, entre otros; que se emplean para fabricar diversos productos como: espumas, adhesivos, recubrimientos, plásticos, textiles, pinturas, solventes, removedores de pinturas, productos de silicona, explosivos, herbicidas y pesticidas [1]. El metanol también se aprovecha como sustituto de los compuestos clorofluorocarbonado en los aerosoles, como limpiador de parabrisas para automóviles y como anticongelante de tuberías [2]. Además, en el sector de los carburantes, el metanol es un componente en la producción de un aditivo para mejorar la combustión de combustibles sin plomo y se usa en la elaboración de biocombustibles, como el biodiesel [3]. Las diferentes aplicaciones mencionadas de metanol demuestran que en las últimas décadas se incrementa su uso y que el porcentaje de eficiencia de los diferentes procesos de transformación no son del 100 %, lo que ocasiona pérdidas de metanol o sus derivados como trazas en el agua residual industrial, lo cual representa un impacto ambiental, puesto que el metanol es considerado un desecho peligroso y no debe descargarse nunca directamente en alcantarillas o aguas superficiales [4], [5]. Sin embargo, en Colombia no se exige en la normatividad [6]–[8] su medición como control en los vertimientos para ser introducidos en las aguas superficiales o subterráneas, interiores o marinas, confirmando la importancia de esta investigación preliminar al evaluar este tipo de compuestos contribuyendo a un conocimiento

más definido sobre la presencia del metanol en residuos líquidos para ejercer control o establecer si su presencia en las aguas residuales genera algún tipo de contaminación con consecuencias a las poblaciones y el medio ambiente, porque de acuerdo al tercer informe bienal de Cambio Climático de Colombia, un 8 % de las emisiones de  $CO_2$  proviene de la disposición de los residuos sólidos y líquidos [9]–[11].

Debido a los diferentes usos que se le ha dado al metanol en los últimos años y de la poca información sobre su tratamiento al ser vertido a corrientes de agua superficial, lleva a pensar que muchas de las industrias no cuentan con sistemas de tratamiento de aguas para este compuesto por no ser exigido, lo que hace que el riesgo de ser vertido en los cuerpos de agua aumente; no obstante, en Colombia se hace un seguimiento a los STAR (Sistemas de Tratamiento de Aguas Residuales), pero la información reportada es general de si existe o no un tratamiento a este tipo de vertimientos [12], donde solo el 33 % (336) de 1008 municipios informan sobre la Planta de Tratamiento de Aguas Residuales (PTAR) y de estos solo el 23 % (84) tienen PTAR para mitigar la influencia de los vertimientos residuales en el ambiente, pero no hay datos que referencien la evaluación de la presencia de metanol [13].

Los problemas ambientales que el metanol puede causar por ser un compuesto tóxico, fácilmente miscible en el agua, que a concentraciones elevadas llegue a afectar los ecosistemas acuáticos, sin ser ajeno este problema al suelo y al aire, causando muerte a: microorganismos, plantas, animales y seres humanos que allí habitan o que de allí se alimentan [5]. Adicionalmente, la US Environmental Protection Agency (EPA) determinó que a concentraciones mayores a 10.000 mg/L, el metanol es tóxico para organismos acuáticos como peces, dafnias y algas. Al ingresar el metanol en ecosistemas acuáticos se afecta significativamente la vida de diferentes organismos y surgen problemas asociados con hiperactividad y convulsión, además de sofocación, secreción mucosa, afectaciones tanto branquiales como a su sistema nervioso central, aumento de la cantidad de dióxido de carbono en la sangre y reducción de la fecundidad [14]. En los seres humanos las concentraciones elevadas el metanol puede causar dolor de cabeza, mareo, náuseas, vómitos y muerte. Una exposición aguda puede causar ceguera o pérdida de la visión, ya que puede dañar seriamente el nervio óptico, además una exposición crónica puede causar daños al hígado o de cirrosis [4].

En la actualidad se conocen diferentes tratamientos de las aguas residuales contaminadas con metanol, que buscan disminuir su concentración, como el carbón activado, la oxidación avanzada, ciclos biológicos, con membranas y biológico con filtros activados [5], [15]–[21]. Los porcentajes de eficiencia de remoción de metanol en sistema aerobio ha sido reportado en un 95 % usando microorganismos metilotróficos encapsulados en polímeros (bio-filtros) [22] y en sistemas anaerobios el uso de metanol; como nutriente para aumentar la eficiencia de reducción de elementos químicos orgánicos ha sido reportado en un 65 % [21], pero no han sido dados a conocer porcentajes de eficiencia referenciada a la eliminación de metanol en lodos activados por oxidación como el presentado en este artículo. Los mecanismos de remoción biológica se basan en la oxidación del metanol [22] lo cual se evidencia en un sistema de lodos activados aerobio o anaeróbicamente a través de la metanogénesis [23]. El uso de estas metodologías busca reducir el impacto ambiental de la contaminación presente en aguas residuales.

Lo anterior plantea un claro panorama frente a los riesgos que representa el metanol en térmi-

nos de impactos ambientales, pero pocos son los datos que se encuentran sobre el uso de sistemas convencionales como los lodos activados, a partir de desechos industriales, pues es una alternativa biotecnológica en el tratamiento de aguas contaminadas con compuestos complejos como el metanol [24]. Los lodos activados son lodos sedimentados de las aguas residuales crudas previamente agitados en la presencia de abundante oxígeno atmosférico [25]. Un sistema de lodos activados emplea una suspensión de microorganismos, como bacterias, hongos, protozoos y rotífero, los cuales se encargan de disolver compuestos carbonados, más comúnmente conocido como materia orgánica, convirtiéndola en productos más simples y adicionalmente biomasa [24]. Por lo tanto, se propuso como objetivo evaluar los parámetros fisicoquímicos en un reactor discontinuo de lodos activados a escala de laboratorio para el tratamiento de un agua de origen industrial contaminada con metanol, con el propósito de establecer si hay diferencias en los parámetros de DBO, DQO y SS durante el tratamiento del compuesto químico.

## 2. Materiales y métodos

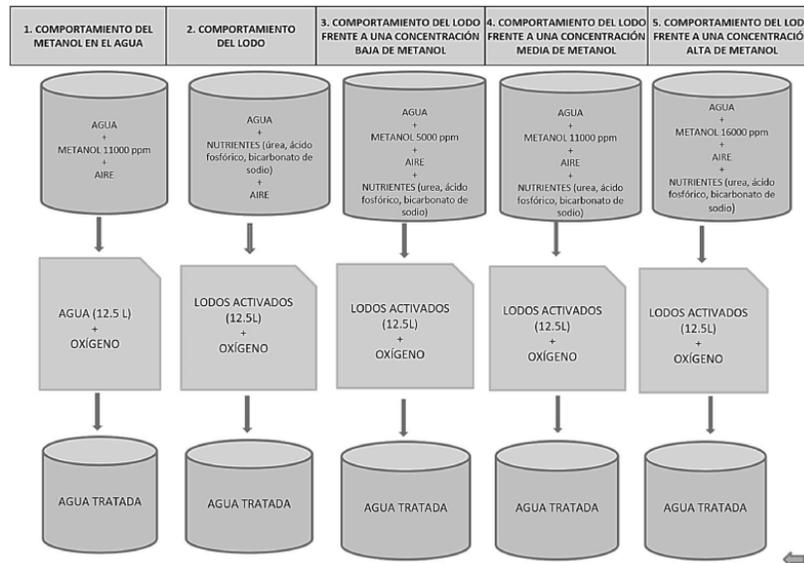
### 2.1. Lodo activado

Este material para la experimentación fue obtenido del zanjón de oxidación de la planta de tratamiento de aguas residuales de la empresa Alpina Productos Alimenticios S.A. en su sede de Facatativá (Departamento de Cundinamarca). Se caracterizó los lodos y se estableció que era de color café claro, con olor similar al de tierra húmeda, con pH de 7, con una demanda química de oxígeno (DQO) de 3.680 mg/L y sólidos suspendidos de 5.403,3 mg/L. Se decidió trabajar con este tipo de lodo por tener una alta cantidad de nutrientes y microorganismos metilotróficos y ácidos lácticos, ya que en dicha empresa se producen alimentos que inducen en los microorganismos a la formación de metanol como compuesto organoléptico y su presencia en los lodos les permite a otros organismos su adaptación biológica.

### 2.2. Sistemas de tratamiento discontinuo

A escala de laboratorio se diseñaron y se construyeron en material acrílico de 0.4mm de espesor cinco sistemas o reactores de tratamiento discontinuo (entrada, proceso, salida) de lodos activados cada uno con dos compartimientos: uno para mezcla y otro para la sedimentación y se denominaron así: sistema como blanco o sistema 1 que contenía solamente agua y metanol a 11.000 ppm; sistema como control o sistema 2 que contenía solo los lodos recolectados; sistema con lodos y concentración de metanol de 5.000 ppm o sistema 3; sistema con lodos y concentración de metanol de 11.000 ppm o sistema 4; y sistema con lodos y concentración de metanol de 16.000 ppm o sistema 5. Las concentraciones de metanol de los sistemas 3, 4 y 5 representaban valores bajo, medio y alto con respecto al valor de referencia definido por la EPA [4], es decir, a concentraciones mayores de 10.000 mg/L de metanol es tóxico para organismos acuáticos y a concentraciones mayores de 80mg/100mL en la sangre es perjudicial para seres humanos.

La alimentación de los lodos se realizó por gravedad teniendo presente la relación DQO:N:P = 100:5:1 para garantizar la vida microbiana y depuración del agua con metanol, la aireación se garantizó con piedras difusoras de aire conectadas a un compresor con potencia de 200 psi y se regulo



**Figura 1.** Descripción de los sistemas del montaje experimental para la evaluación de los factores físicoquímicos en la remoción de metanol en agua

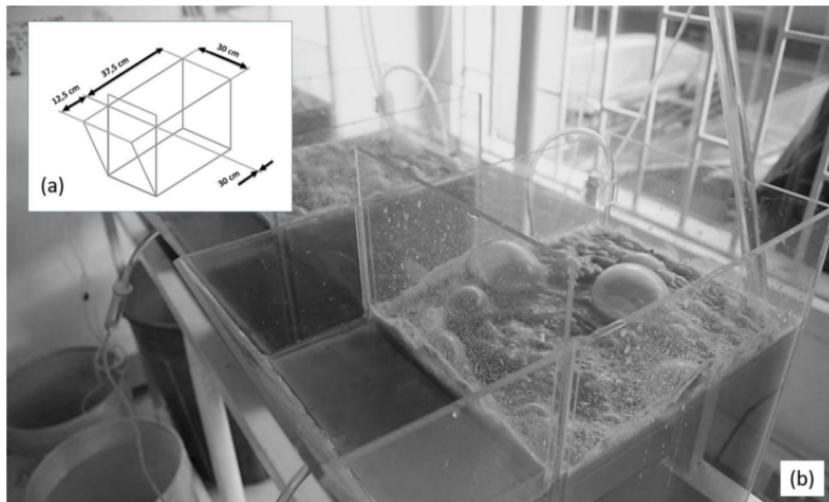
la entrada a 17 psi con un caudal de aire de  $1,73 \times 10^{-6} \text{ m}^3/\text{s}$  (6,23 L/h) y el tiempo de retención de los tratamientos fue de 48 horas, un volumen de 12.5 L de lodos activados y se trabajó el 80 % del mismo (Figura 1 y tablas II y III).

Una vez definidas las particularidades de cada sistema de tratamiento, se procedió a su implementación y evaluación mediante el desarrollo de tres fases que se describen a continuación.

### 2.3. Fase I: diseño e instalación del montaje experimental

Los cinco reactores se diseñaron con dos compartimientos, uno de mezcla y otro para la sedimentación, con las medidas mostradas en la Figura 1 y se construyeron en material acrílico de 0,04 cm de calibre. El compartimiento de mezcla contó con un volumen útil de  $0,010 \text{ m}^3$  (10 L) y el compartimiento sedimentador fue diseñado con un volumen útil de  $0,0025 \text{ m}^3$  (2,5 L), siguiendo el principio del cono de Imhoff [24], [26] con las siguientes dimensiones: ancho/largo/alto 20cm/37,5cm/30cm (figura 2).

Para el suministro de los nutrientes (urea y ácido fosfórico) a los lodos activados en relación de 100:5:1, se realizó diariamente según la concentración de DQO de trabajo, y por gravedad se suministraron los nutrientes desde tanques de alimentación  $0,030 \text{ m}^3$  (30 L) aireados para evitar su sedimentación. A su salida se conectó una válvula y manguera plástica de 8 mm de diámetro nominal, cuya función principal era la regulación del caudal a tratar dentro de los reactores. Además, en los tanques de alimentación se adicionaron agua de acueducto y metanol industrial en las concentraciones fijadas para cada sistema de tratamiento.



**Figura 2.** Reactores discontinuos para tratar lodos activados a escala de laboratorio. (a) Medidas del biorreactor aerobio y (b) montaje experimental del reactor con lodos activados.

**Fecha de captura:** 07/11/2013. **Formato de archivo:** JPG. **Cámara:** NIKON D90.

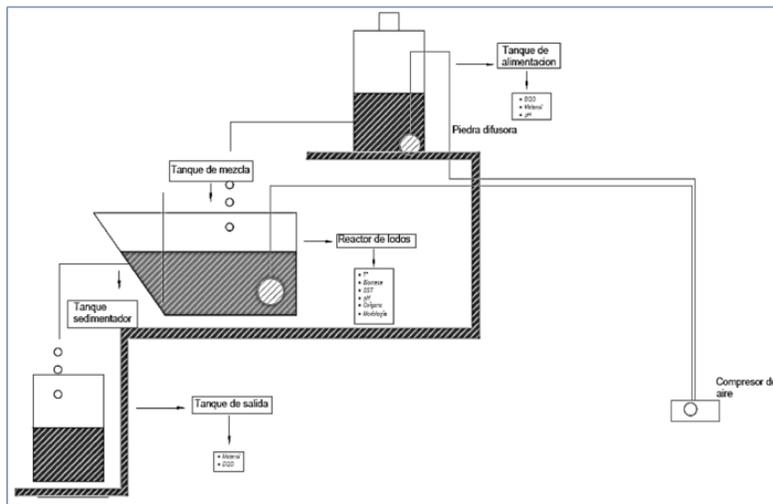
Todas las partes del proceso (alimentación, tratamiento, y recolección del agua tratada) fueron ubicadas de forma escalonada en un soporte metálico de 1,55 m de altura, con lo cual se buscó hacer uso de la fuerza de gravedad y aprovechar la caída en cada uno de los puntos de tratamiento, para garantizar un flujo constante tanto del agua a tratar como del agua tratada (Figura 3).

## 2.4. Fase II: rutina operativa

**Tabla I.** Parámetros de control de los sistemas de tratamientos experimentales

Parámetros	Métodos	Ubicación de la toma de muestra
Oxígeno disuelto	Oximetría [27] digital con electrodo marca WTW	Reactor
Sólidos solubles totales	Secado en estufa [27]	Reactor
Demanda química de oxígeno (DQO)	Solución de digestión [28]	Tanque de alimentación y tanque de salida
Contenido de metanol	Colorimétrico [29]	Tanque de alimentación y tanque de salida
Temperatura	Lectura directa [27] en el oxímetro digital con electrodo marca WTW	Reactor
pH	Papel indicador [27] con intervalo de pH modelo CVQ2055	Reactor

Luego del montaje de los cinco sistemas de tratamientos, se llevaron a cabo los controles presentados en la Tabla I a las 0, 24 y 48 h con tomas de muestras para los análisis fisicoquímicos con sus respectivos métodos por cuadruplicado. Cada uno de los parámetros medidos se realizó bajo el mismo protocolo establecido todos los días y los datos eran consignados en una tabla diseñada para la tabulación de los datos por sistema de tratamiento. Además, cada sistema de tratamiento fue alimentado a diario, asegurándose siempre un pH neutro a los nutrientes almacenado en los tanques de alimentación, los cuales conducían posteriormente al suministro de los mismos a cada uno de los reactores. La determinación de DQO y contenido de metanol se trabajó con las normas del Instituto Colombiano de Normas Técnicas y Certificación (ICONTEC), pues es el organismo nacional encargado de la normalización de métodos.



**Figura 3.** Distribución escalonada de los equipos implementados a escala de laboratorio para cada sistema de tratamiento de lodos activados con el tanque de alimentación en la parte superior, seguido del reactor de lodos en la parte media y finalmente el tanque de salida en la parte inferior y la distribución del aire desde el compresor al tanque de alimentación y al reactor

### 2.5. Fase III: cálculos físicoquímicos

Con los resultados de los análisis físicoquímicos de la Fase II, fue hallada la pérdida de metanol de cada sistema de tratamiento en cada repetición y se obtuvo un valor promedio para determinar si el porcentaje de remoción de metanol fue bajo o alto y, de esta manera, evaluar la eficiencia de los lodos activados en cada sistema de tratamiento.

### 2.6. Evaluación estadística

Como modelo estadístico se aplicó un análisis de varianza (ANOVA) mixto univariado para un diseño experimental de medidas repetidas en el programa SAS versión 9.1.3 ( $p \leq 0,05$ ). Este modelo se aplicó a los resultados de demanda química de oxígeno, de concentración de metanol y de sólidos solubles totales, para observar el comportamiento en general entre los sistemas de tratamiento, el tiempo y la relación tiempo-sistema. Los datos se tomaron por cuadruplicado y se realizó tres veces este procedimiento en cada sistema.

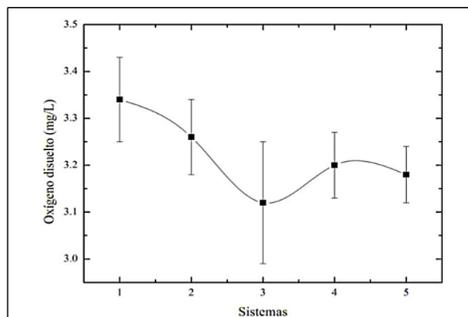
**Tabla II.** Componentes de los tanques de alimentación para cada sistema de tratamiento

Componentes	Reactor 1	Reactor 2	Reactor 3	Reactor 4	Reactor 5
Agua (L)	68,5	68,5	68,5	68,5	68,5
Metanol (mL)	865,1	No aplica	432,6	865,1	1297,7
Urea (g)	No aplica	12,0	17,1	34,3	51,4
Ácido fosfórico (mL)	No aplica	2,4	3,4	6,9	10,3
Bicarbonato de sodio (g)	No aplica	11,0	17,0	34,0	48,0
pH	6	6	7	7	7

**Tabla III.** Parámetros utilizados en el diseño del reactor de lodos activados

Parámetro	Valor	Unidad de medida
Volumen total del reactor de mezcla completa	12.5	L
Volumen del tanque de aireación	10	L
Volumen del sedimentador	2.5	L
Caudal de entrada	1.43	L/h
Caudal de salida	1.43	L/h
Tiempo de retención hidráulico	7	h

En la Tabla II se muestran las cantidades con las que fueron operados cada tanque de alimentación durante la etapa experimental para los cinco sistemas de tratamiento y en la Tabla III los seis parámetros que se tuvieron en cuenta para el diseño del reactor con lodos activados.

**Figura 4.** Comportamiento del oxígeno disuelto en los cinco sistemas de tratamiento de lodos activos para biodegradación del metanol.

### 3. Resultados y discusión

#### 3.1. Oxígeno disuelto, temperatura y pH

El parámetro fisicoquímico de oxígeno disuelto fue estable a través del tiempo, ya que su valor promedio siempre estuvo por encima de 3mg/ L para todos los sistemas de tratamiento (Figura 4), con lo cual se garantizó que cada uno de ellos contara con la cantidad de oxígeno suficiente para que se llevaran a cabo los procesos de biodegradación del metanol (sistemas 2, 3, 4 y 5); sin ser la excepción el sistema 1, el cual contó con aireación con el fin de garantizar las mismas condiciones para todos los tanques.

En cuanto a los valores de temperatura oscilaron en un rango bastante bajo durante la etapa experimental (16-17°C) en cada uno de los sistemas. El parámetro de pH no presentó mucha variación durante la etapa experimental en ninguno de los sistemas durante las repeticiones (6-7,5); este hecho garantizó a los diferentes reactores condiciones de pH neutro o de valores cercanos.

#### 3.2. Sólidos solubles totales, demanda química de oxígeno y contenido de metanol

El método empleado para la medición de metanol determinó una pérdida promedio en el sistema 1 del 11,3 %, valor que corresponde a la volatilización de este compuesto debido a que no contaba con sólidos solubles. El comportamiento de los datos arrojados por el sistema 2, que no contaba con ingreso de metanol, mostró un comportamiento estable en la cantidad de sólidos suspendidos totales en todas las repeticiones, hecho que no es ajeno a los resultados arrojados por la demanda química de oxígeno, manifestando buenas condiciones del lodo activado para ser usado como tratamiento biológico.

Los datos arrojados para el sistema 3, presentaron que el comportamiento de la demanda química de oxígeno no fue homogéneo, se observó una clara tendencia a través del tiempo a aumentar por encima del valor inicial (5.120 ppm), debido a los problemas de sedimentabilidad del lodo en el tanque sedimentador, dejando partículas del lodo en suspensión, las cuales salieron del sistema ocasionando que la cantidad de materia orgánica aumentara la demanda química de oxígeno y disminuyera la cantidad de sólidos. Los datos para la medición del metanol presentaron una clara disminución de la concentración, con un promedio de 11,7 % de pérdidas (figuras 5 y 6).

En el sistema 4, fue evidente que los valores arrojados por cada una de las réplicas para la demanda química de oxígeno no tienen un comportamiento estable, debido a que los valores aumentaban y disminuían. Por otra parte, la medición de metanol muestra un promedio de pérdida del 12,9 % mientras que los sólidos suspendidos disminuyeron significativamente en todas las repeticiones (Figuras 5 y 6).

El sistema 5 presentó disminución en cuanto al parámetro de la demanda química de oxígeno hasta la mitad de la experimentación y aumentó en la otra mitad, mientras que la concentración de metanol en el sistema indicó un promedio de pérdidas del 5,3 %. Caso similar sucedió con la cantidad de sólidos en el sistema los cuales también demostraron una clara tendencia a la baja para

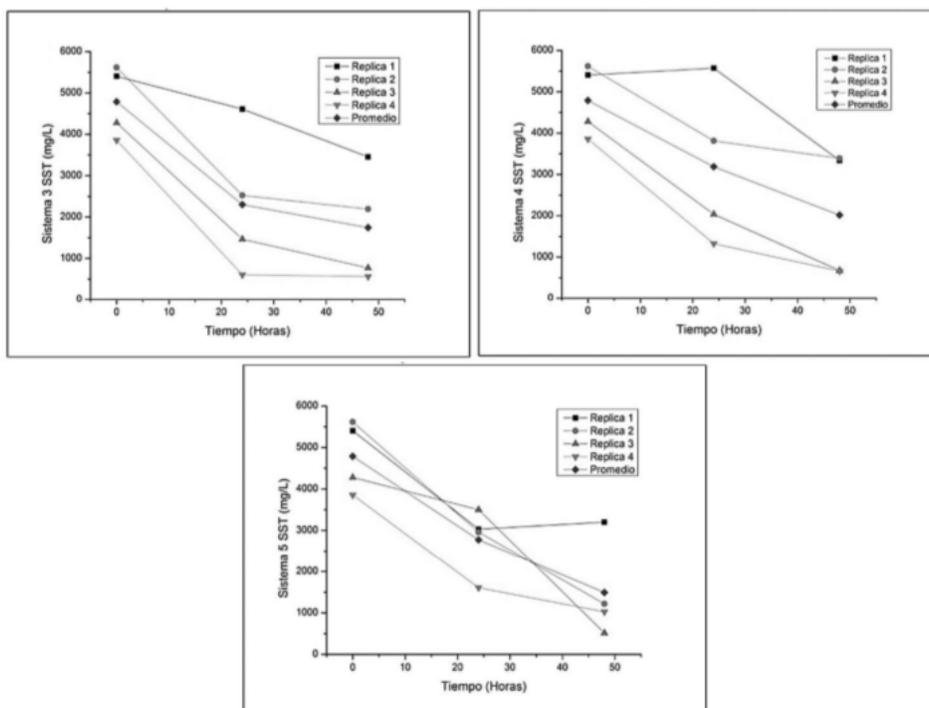


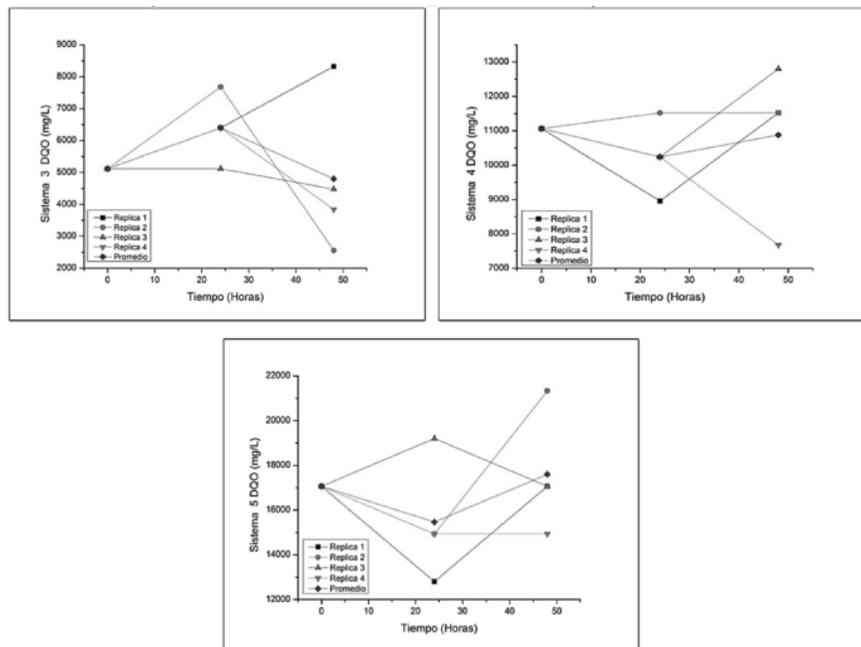
Figura 5. Comportamiento decreciente de los promedios de sólidos solubles totales en el tiempo para los reactores que manejaban lodos activados

todas las repeticiones (figuras 5 y 6).

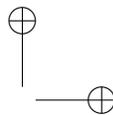
Por consiguiente, la pérdida promedio de metanol obtenida para la experimentación en los sistemas de tratamiento 3, 4 y 5 por lodos activos, fue del 10,3 % dejando un contenido de metanol remanente de 89,7 %. El resultado anterior demostró que el tratamiento de lodos activos para la biodegradación de metanol en aguas residuales industriales es posible puesto que se han realizado otros estudios al tratar aguas residuales con un contenido de metanol menor del 10 % y en 24 h se logra disminuir el contenido de metanol [30]. Además, el control del metanol es importante pues estudios demuestran diversas intoxicaciones por este componente [31].

### 3.3. Análisis estadístico

Se analizaron solo tres parámetros (DQO, SST y concentración de metanol) a través del modelo estadístico y en todos los análisis de varianza el p-valor fue menor que 0,05, rechazándose todas las hipótesis nulas propuestas en la investigación y permite definir que: a) Para el parámetro DQO: las curvas de decrecimiento de DQO para los cinco sistemas no son paralelas, que si hay decrecimiento de DQO a lo largo de los tres tiempos promedio de medición de los cutriplicados y que hay diferencia entre los cinco sistemas; b) Para el parámetro SST: las curvas de decrecimiento de SST para los cinco sistemas no son paralelas, que si hay decrecimiento de SST a lo largo de los tres tiempos de medición promedio de medición de los cutriplicados y que hay diferencia entre los



**Figura 6.** Comportamientos de los promedios de demanda química de oxígeno variable y contenido de metanol decreciente en el tiempo para los reactores que manejaban lodos activados



cinco sistemas para el parámetro concentración de metanol (CM): las curvas de decrecimiento de la CM para los cinco sistemas no son paralelas, que si hay decrecimiento de la CM a lo largo de los tres tiempos de medición promedio de medición de los cutriplificados y que hay diferencia entre los cinco sistemas.

## 4. Conclusiones

Al analizar el comportamiento del metanol en el sistema 1 se evidenció pérdidas significativas de metanol (11,2%), lo que confirmó su propiedad de volatilización [5]. Por su parte el sistema 2 mantuvo un comportamiento estable en cuanto a la cantidad de sólidos suspendidos, lo cual significó que las pérdidas del mismo fueron pocas con respecto a los otros sistemas, demostrando que el lodo cumplía con los requisitos óptimos de alimentación, pH y oxigenación para tener un buen funcionamiento como tratamiento biológico de aguas residuales.

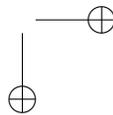
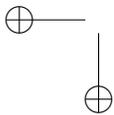
En la etapa operativa de los sistemas de tratamiento 3, 4 y 5 presentaron variaciones en diferentes análisis, entre ellos la demanda química de oxígeno, la cual no tuvo un comportamiento esperado debido al bajo índice de sedimentabilidad del lodo, ya que no fue posible observar una separación en dos fases (agua clarificada y lodo sedimentado), sino que por el contrario se evidenciaron partículas de lodo que al no sedimentar cedieron a la hidráulica del sistema y salieron del medio junto con el agua tratada, lo cual ocasionó que la cantidad de sólidos suspendidos totales disminuyera en cada uno de los sistemas, y hubiera un aumento en la demanda química de oxígeno.

Debido a la poca información existente sobre el comportamiento de los lodos activados como sistema biológico frente a un sustrato contaminante como el metanol, los resultados de esta investigación son significativos, al ser de manera parcial los primeros en pretender conocer la degradación de este compuesto, teniendo en cuenta algunos parámetros físicoquímicos y abre el camino para estudiar como el metanol influye en la cinética de reacción de la biomasa presente en los lodos para lograr procesos de remoción más eficiente, evaluar el efecto de diferentes tiempos de estabilización de los Lodos, la evaluación de otros sistemas como los anaeróbicos y en posteriores trabajos se pretende determinar su viabilidad técnica y económica, después de establecer su eficiencia química y biológica en la remoción de metanol a partir de aguas residuales industriales.

Con los datos obtenidos se podrán generar nuevas propuestas de biorreactores acoplados a este tipo de compuestos químicos para mejorar el nivel de eficiencia en la remoción de metanol y con los datos que se obtengan de la segunda fase del trabajo donde se busca establecer el comportamiento biológico y cinético de los lodos activos, ya que el objetivo de este primer trabajo era establecer si el metanol podía ser oxidado y por lo tanto existiera la posibilidad de ser tratado por medio de este tipo de tecnología biotecnológica ampliamente trabajada para el tratamiento de vertimientos industriales.

## Referencias

- [1] Methanex Corporation, *Formulario de información anual*. 2011 [En línea]. Disponible en: <http://www.methanex.cl/noticias/2011/noticia0324.pdf> ↑99



- [2] Osorio, R., *Metanol obtención y usos*. Petroblogger.com: Blog sobre la industria de petróleo y gas natural. 2011 [En línea]. Disponible en: <http://www.ingenieriadepetroleo.com/metanol-obtencion-y-usos.html> ↑99
- [3] E. Rodríguez y M. Uribe Echavarría, *Obtención de metanol: propiedades-usos*. Editorial de la Universidad Tecnológica Nacional. S.f. [En línea]. Disponible en: [www.edutecne.utn.edu.ar/procesos\\_fisicoquimicos/Obtencion\\_de\\_metanol.pdf](http://www.edutecne.utn.edu.ar/procesos_fisicoquimicos/Obtencion_de_metanol.pdf) ↑99
- [4] Methanol Institute, *Manual de manipulación segura del metanol*. 2012 [En línea]. Disponible en: <http://methanol.org/Health-And-Safety/Safety-Resources/Health—Safety/Methanol-Safe-Handling-Manual-Spanish.aspx> ↑99, 100, 101
- [5] L. Yuan, W. Zhi, Y. Liu, E. Smiley, D. Gallagher, X. Chen, A.M. Dietrich and H. Zhang, “Degradation of cis- and trans-(4-methylcyclohexyl) methanol in activated sludge Journal of hazardous materials”. *Journal of Hazardous Materials*, vol. 306, pp. 247 – 256, 2016. ↑99, 100, 108
- [6] República de Colombia, Artículos 20 y 21: Decreto 1594 de 1984. [En línea]. Disponible en: <http://www.alcaldia bogota.gov.co/sisjur/normas/Norma1.jsp?i=18617> ↑99
- [7] Ministerio de Ambiente, Vivienda y Desarrollo Territorial, Decreto 3930 de 2010. [En línea]. Disponible en: <http://www.alcaldia bogota.gov.co/sisjur/normas/Norma1.jsp?i=40620#> ↑99
- [8] Ministerio de Ambiente y Desarrollo Sostenible, Decreto 303 de 2012 [En línea]. Disponible en: <https://www.crq.gov.co/Documentos/JURI-DICA/decreto%20303%20de%202012.pdf> ↑99
- [9] Ministerio de Medio Ambiente y Desarrollo Sostenible, *Colombia se prepara para su participación en la COP 21 en París*. 2016 [En línea]. Disponible en: <http://www.minambiente.gov.co/index.php/component/content/article/126-noticias-cambio-climatico/2115-colombia-se-prepara-para-su-participacion-en-la-cop-21-en-paris-2> ↑100
- [10] IDEAM-ONU, *Información científica al servicio de la planificación. Tercera comunicación nacional de cambio climático*. 2016 [En línea]. Disponible en: [https://prezi.com/j6gz4upbmi74/informacion-cientifica-al-servicio-de-la-planificacion-sectorial-ideam-onu/?utm\\_campaign=share&utm\\_medium=copy](https://prezi.com/j6gz4upbmi74/informacion-cientifica-al-servicio-de-la-planificacion-sectorial-ideam-onu/?utm_campaign=share&utm_medium=copy) ↑100
- [11] IDEAM, UNDP y Republica de Colombia, *Primer Informe bienal de actualización de Colombia ante la convención marco de las naciones unidas sobre el cambio climático*. 2015 [En línea]. Disponible en: [http://documentacion.ideam.gov.co/openbiblio/bvirtual/023422/Primer\\_Informe\\_Bienal.pdf](http://documentacion.ideam.gov.co/openbiblio/bvirtual/023422/Primer_Informe_Bienal.pdf) ↑100
- [12] Superintendencia de Servicios Públicos, Informe técnico sobre sistemas de tratamiento de aguas residuales en Colombia. 2013 [En línea]. Disponible en: <http://www.superservicios.gov.co/content/download/4989/47298> ↑100
- [13] Procuraduría General de la Nación, “Capítulo 1: El estado del agua, el alcantarillado y los residuos sólidos en los municipios”. *El agua potable y el saneamiento básico en los planes de desarrollo*. S.f. [En línea]. Disponible en: [http://www.procuraduria.gov.co/portal/media/file/descargas/publicaciones/parte3\\_agua.pdf](http://www.procuraduria.gov.co/portal/media/file/descargas/publicaciones/parte3_agua.pdf) ↑100
- [14] A. Kaviraj, F. Bhunia, and N.C. Saha, “Toxicity of Methanol to Fish, Crustacean, Oligochaete Worm, and Aquatic Ecosystem”. *Int J Toxicol*. vol. 23, pp. 55-63. 2004. ↑100
- [15] C. Idrobo, y J. Gutierrez, “Electrooxidación catalítica de metanol en agua”. *Revista Biotecnología*, vol 4. no. 1. pp. 87-94, marzo, 2006. [En línea]. Disponible en: <http://revistabiotecnologia.unicauca.edu.co/revista/index.php/biotecnologia/article/viewFile/32/21> ↑100
- [16] N.R. Louzeiro, D.S. Mavinic, W.K. Oldham, A. Meisen and I.S. Gardner, “Methanol-induced biological nutrient removal kinetics in a full-scale sequencing batch reactor”. *Water research*. vol. 36, pp. 2721–2732, 2002. ↑100
- [17] L. Narváez, R. Cáceres y O. Marfà, “Depuración de lixiviados procedentes de viveros de ornamentales mediante humedales artificiales. Eficacia del metanol como fuente carbonada”. En *VI Jornadas Ibéricas de Horticultura Ornamental*, Valencia, España, 1-3 octubre. Actas de horticultura. no. 68. pp. 239–244, 2014. ↑100
- [18] B. Liu and B. Li, “Single chamber microbial fuel cells (SCMFCs) treating wastewater containing methanol”. *International journal of hydrogen energy*, vol.39, pp. 2340-2344. 2014 ↑100
- [19] G. Cao, Y. Zhang, L. Chen, J. Liu, K. Mao, K. Li and J. Zhou, “Production of a biofloculant from methanol wastewater and its application in arsenite removal”. *Chemosphere*, vol. 141, pp. 274-281, 2015. ↑100
- [20] E. Shimoda, T. Fujii, R. Hayashi and Y. Oshima, “Kinetic analysis of the mixture effect in supercritical water oxidation of ammonia/methanol”. *J. of supercritical fluids*, vol. 116, pp. 232-238, 2016. ↑100
- [21] E. Torresi, M. Escolà Casas, F. Polesel, B.G. Plósz, M. Christensson and K. Bester, “Impact of external carbon dose on the removal of micropollutants using methanol and ethanol in post-denitrifying moving bed biofilm reactors”. [En línea] *Water Research*. In Press, Accepted Manuscript, Oct, 2016, doi:10.1016/j.watres.2016.10.068 ↑100
- [22] O.A. Kamanina, D.G. Lavrova, V.A. Arlyapov, V.A. Alferov, and O. N. Ponamoreva, “Silica sol-gel encapsulated methylotrophic yeast as filling of biofilters for the removal of methanol from industrial wastewater”. *Enzyme and*

- Microbial Technology*, vol. 92, pp. 94-98, 2016. ↑100
- [23] S. Saha, N. Badhe, J. De Vrieze, R. Biswas and T. Nandy, "Methanol induces low temperature resilient methanogens and improves methane generation from domestic wastewater at low to moderate temperatures". *Bioresource Technology*, vol. 189, pp. 370-378, 2015. ↑100
- [24] Metcalf and Eddy, *Wastewater Engineering Treatment and Reuse*. Fifth Edition. New York, Mc Graw Hill Education, 2014. ↑101, 102
- [25] T. Matsumoto and I. Sánchez Ortiz. "Desempeño de la planta de tratamiento de aguas residuales de São João de Itacema (Brasil)". *Revista Ingeniería*, vol.21, no. 2, pp. 176-186, abril, 2016. [En línea]. Disponible en: <http://revistas.udistrital.edu.co/ojs/index.php/reving/article/view/9970/11563> ↑101
- [26] Eddy, & Metcalf, *Procesos biológicos unitarios*. Universidad de Salamanca. Centro de investigación de desarrollo tecnológico del agua. 1996. [En línea]. Disponible en: [http://cida.usal.es/residuales/libros/logo/pdf/Procesos\\_biologicos\\_unitarios.pdf](http://cida.usal.es/residuales/libros/logo/pdf/Procesos_biologicos_unitarios.pdf) ↑102
- [27] F.J. Molina Perez and D.C. Rodríguez, *Manual de laboratorio de procesos biológicos*. Universidad Nacional de Colombia, Medellín, Colombia. 2008. ↑103
- [28] ICONTEC, *Norma técnica colombiana 3629. Calidad del agua.demanda química de oxígeno (DQO)*. ICONTEC, Bogotá D.C., 2002. ↑103
- [29] ICONTEC, *Norma Técnica Colombiana 5159. Bebidas alcoholicas. Método colorimétrico para determinar el contenido de metanol*. ICONTEC, Bogotá D.C., 2003. ↑103
- [30] Al-D. Salam, "Methanol removal from methanol-water mixture using municipal activated sludge". *Journal of Engineering Science and Technology*, vol 8, no 5, pp 578-587, 2013. ↑107
- [31] Instituto Nacional de Salud, *Protocolo de vigilancia y control de intoxicaciones por metanol*. Bogotá D.C., 2011. ↑107

### Marvin Caravali

Nació en Puerto Tejada (Cauca), Colombia. Es Ingeniero Ambiental de la Universidad Distrital Francisco José de Caldas, de Bogotá, Colombia. Se desempeñó como estudiante en la Universidad Distrital durante 5 años, donde desarrolló el proyecto código: 3-180-328-12 financiado por el CIDC (Centro de investigaciones y desarrollo científico) de la Universidad Distrital.

Correo electrónico: marvin211986@yahoo.com

### Martha L. Lote

Nació en Bogotá, Colombia. Es Ingeniera Ambiental de la Universidad Distrital Francisco José de Caldas, de Bogotá, Colombia. Se desempeñó como estudiante en la Universidad Distrital durante cinco años, donde desarrolló el proyecto código: 3-180-328-12 financiado por el CIDC (Centro de investigaciones y desarrollo científico) de la Universidad Distrital.

Correo electrónico: malote2503@yahoo.es

### Lena Carolina Echeverry

Nació en Bogotá, Colombia. Microbióloga Industrial de la Pontificia Universidad Javeriana, de Bogotá, Colombia. Obtuvo su título de Maestría en Ciencias Biológicas en la Pontificia Universidad Javeriana de Bogotá, Colombia. Actualmente se desempeña como profesora en el área de microbiología en la Universidad Distrital Francisco José de Caldas de Bogotá, Colombia, y pertenece como investigadora al grupo GIA-UD (Grupo de investigación de Ingeniería Ambiental de la Universidad Distrital) donde realiza estudios sobre microbiología, biotecnología y tecnologías apropiadas. Correo electrónico: lcecheverry@udistrital.edu.co

## Methodology

**Detection of Outliers and Imputing of Missing Values for Water Quality UV-VIS Absorbance Time Series***Detección de Valores Extremos e Imputación de Valores Faltantes para la Calidad de Agua en Series de Tiempo de Absorbancia UV-VIS***Leonardo Plazas-Nossa<sup>\*1</sup>, Miguel A. Ávila A.<sup>1</sup>, Andres Torres<sup>2</sup>**<sup>1</sup> Universidad Distrital Francisco José de Caldas. Bogotá - Colombia,<sup>2</sup> Pontificia Universidad Javeriana. Bogotá - Colombia\*Correspondence: [lplazasn@udistrital.edu.co](mailto:lplazasn@udistrital.edu.co)

Recibido: 05-04-2016. Modificado: 19-09-2016. Aceptado: 03-01-2017

## Abstract

**Context:** The UV-Vis absorbance collection using online optical captors for water quality detection may yield outliers and/or missing values. Therefore, pre-processing to correct these anomalies is required to improve the analysis of monitoring data. The aim of this study is to propose a method to detect outliers as well as to fill-in the gaps in time series.

**Method:** Outliers are detected using Winsorizing procedure and the application of the Discrete Fourier Transform (DFT) and the Inverse of Fast Fourier Transform (IFFT) to complete the time series. Together, these tools were used to analyse a case study comprising three sites in Colombia ( (i) Bogotá D.C. Salitre-WWTP (Waste Water Treatment Plant), influent; (ii) Bogotá D.C. Gibraltar Pumping Station (GPS); and, (iii) Itagüí, San Fernando-WWTP, influent (Medellín metropolitan area) ) analysed via UV-Vis (Ultraviolet and Visible) spectra.

**Results:** Outlier detection with the proposed method obtained promising results when window parameter values are small and self-similar, despite that the three time series exhibited different sizes and behaviours. The DFT allowed to process different length gaps having missing values. To assess the validity of the proposed method, continuous subsets (a section) of the absorbance time series without outlier or missing values were removed from the original time series obtaining an average 12 % error rate in the three testing time series.

**Conclusions:** The application of the DFT and the IFFT using the 10 % most important harmonics of useful values, can be advantageous for its later use in different applications, specifically for time series of water quality and quantity in urban sewer systems. One potential application would be the analysis of dry weather affecting rainy seasons, a feature achieved by detecting values that correspond to unusual behaviour in a time series. Additionally, the results hint at the potential of the method in correcting other hydrologic time series.

**Keywords:** Imputing missing values, outlier detection, UV-vis absorbance, water quality, winsorizing.

**Language:** English



Citación: L. Plazas, M. A. Ávila, A. Torres, "Detection of Outliers and Imputing of Missing Values for Water Quality UV-VIS Absorbance Time Series" INGENIERÍA, vol. 22, no. 1, pp. 111-124, 2017.

©Los autores; titular de derechos de reproducción Universidad Distrital Francisco José de Caldas. En línea DOI: <http://dx.doi.org/10.14483/udistrital.jour.reveng.2017.1.a07>

### Resumen

**Contexto:** El registro de la absorbancia UV-Vis mediante captores ópticos en línea para la detección de la calidad del agua, en donde se pueden presentar valores atípicos o valores faltantes. Por lo tanto, el pre-procesamiento para corregir dichas anomalías es necesario para un mejor análisis de los datos de monitoreo. El objetivo de este estudio es proponer un método para detectar e imputar valores extremos como también completar valores faltantes en series de tiempo.

**Método:** La detección de valores atípicos utiliza el procedimiento de enventaneo y la aplicación de la Transformada Discreta de Fourier (DFT –Discrete Fourier Transform) y la inversa de la Transformada Rápida de Fourier (IFFT–Inverse of Fast Fourier Transform) para completar las series de tiempo. Estas herramientas fueron utilizadas para un caso de estudio compuesto por tres sitios en Colombia (i) PTAR-Salitre (Planta de Tratamiento de Aguas Residuales) Bogotá D.C., afluyente; (ii) Estación Elevadora de Gibraltar Bogotá D.C.; y (iii) PTAR-San Fernando, área metropolitana de Medellín, afluyente) analizados mediante espectros UV-Vis (Ultravioleta y Visible).

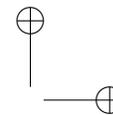
**Resultados:** La detección de valores atípicos con el método propuesto obtiene resultados prometedores cuando los valores de los parámetros de la ventana son pequeños y auto-similares, esto a pesar de que las tres series de tiempo utilizadas presentan diferentes tamaños y comportamientos. Para validar la metodología propuesta, sub-conjuntos continuos (una sección) de las series de tiempo de absorbancia sin valores ausentes o atípicos, fueron removidos de las series original obteniéndose tasas de error de 12 % en promedio para todos los tres sitios de estudio.

**Conclusiones:** La aplicación de la DFT y la IFFT, utilizando el 10 % de los armónicos más importantes de los valores útiles es crucial para su posterior uso en diferentes aplicaciones, específicamente para series de tiempo de calidad y cantidad de agua en sistema de saneamiento urbano. Una posible aplicación podría ser la comparación de los efectos de clima seco respecto a temporadas de lluvia, mediante la detección de valores que corresponden a comportamiento inusual en una serie de tiempo. Además, los resultados indican potencial aplicación futura en la corrección de otras series de tiempo hidrológicas.

**Palabras clave:** Absorbancia UV-Vis, calidad de agua, detección de valores extremos, enventaneo, imputación de valores faltantes.

## 1. Introducción

Continuous on-line measurements such as UV-Vis spectrometry is increasingly applied technique for water quality measurement in sewer systems [1]–[5]. These continuous time series [6]–[8] help to estimate pollutant concentrations in sewer systems and offer real-time control applicability [9]–[13]. Absorbance (essentially surrogate measurements for TSS – Total Suspended Solids or COD – Chemical Oxygen Demand) time series can be used to estimate the dry weather contribution to total TSS and COD loads measured during storm events [14], [15]. In fact, most models of storm weather pollutant loads in combined sewer systems are based on the assumption that total storm event load is the sum of dry and wet weather contributions, with the latter including surface runoff plus possible erosion of deposits accumulated in sewers [16]–[19]. Therefore, during rain events, it is important to distinguish the fraction of flow rate corresponding to dry weather and that corresponding to wet weather. This process, however, is hindered by the fact that the time series may present outliers. Johnson and Wichern [20] define an outlier as “an observation in a data set which appears to be inconsistent with the remainder of that set of data” [21]–[23], [25], [28]. Researches



had been used different methodologies to detect the isolated and extremely peak values as PCA by [23], also using statistical procedures based on prior knowledge of the system that produces the data [21], [24]–[28]. Others researches have been applied artificial intelligence and machine learning methodologies [22], [29]–[31]. In addition, lost values, due to obstructions on the sensors themselves or the occasional sensors removal required for maintenance, would be also present; researches have been using different methodologies for infilling missing values in data sets as Artificial Neural Networks (ANN) by [32], Evolutionary Algorithms by [33], Genetics Algorithms by [34], clustering by [35], using PCA by [36], Bayes' theorem [37], linear and non-linear regressions [38]–[40]. Several detection outlier techniques require to obtain a fitted model [41] to assess which are the extremely peaks to be removed and the infilling methods require multivariate information to obtain a model to fill the gaps in data sets.

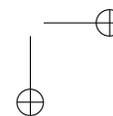
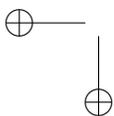
The methodology laid out in this article detects and removes outliers and fill gaps making the missing values imputation in time series. In theory, the proposed procedure combines statistical values, outlier detection [42], [43], DFT and IFFT to complete time series. Thus, it is used the combination of the Winsorising and DFT procedures to perform the task of detection, removal of outliers and imputing of missing values to maintain the observed periodic behaviour of these time series. In practice, this methodology was applied to three time series at the same number of study sites in Colombia for UV-Vis spectra. Locations of the study sites are as follows: (i) Bogotá D.C. Salitre-WWTP (Waste Water Treatment Plant), influent; (ii) Bogotá D.C. Gibraltar Pumping Station (GPS); and, (iii) Itagüí, San Fernando-WWTP, influent (Medellín metropolitan area).

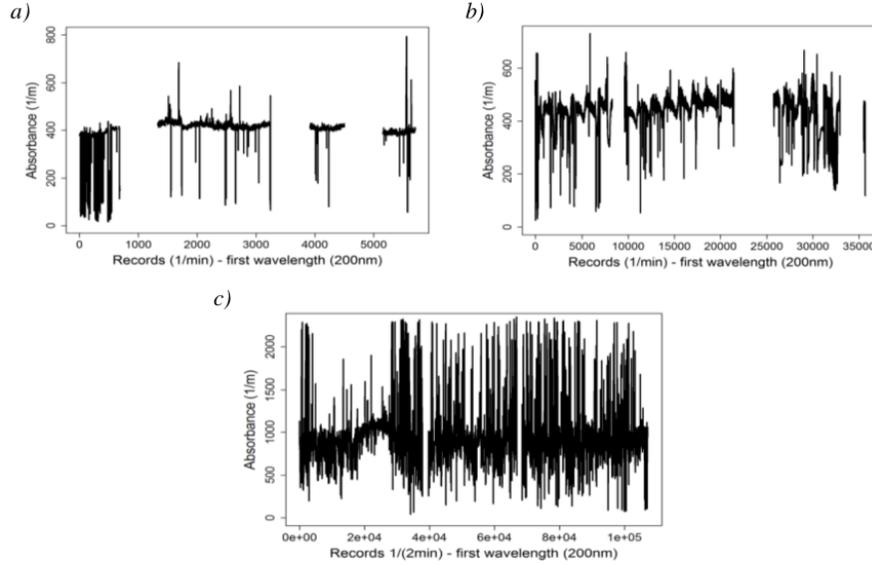
## 2. Materials and Methods

The spectro::lyser<sup>TM</sup> UV-Vis sensors deployed at the three Colombian sites are submersible probes with a length of approximately 65 cm and a diameter of approximately 44 mm. Functionally speaking, they register light attenuation (absorbance) on-line in relative continuous time (one signal per minute). These sensors work with a light source provided by a Xenon lamp for wavelengths ranging from 200 nm to 750 nm, with intervals of 2.5 nm [1], [44].

With regard to the duration of the study, i.e. data gathered, the three time series are displayed in Figure 1: (i) Bogotá D.C. Salitre-WWTP, influent (5705 records, one per minute, from June 29th, 2011 at 9:03 h to July 3rd, 2011 at 17:33 h); (ii) Bogotá D.C. Gibraltar Pumping Station (GPS) (35684 records, one per minute, from October 18th, 2011 at 16:17 h to November 11th, 2011 at 11:20 h); and (iii) Itagüí, San Fernando-WWTP, influent (Medellín) (107204 records, one every two minutes, from September 24th, 2011 at 11:08 h to February 20th, 2011 at 10:18 h.

As Figure 1 aptly demonstrates, outliers accompanied our data collection. Outliers, those data that deviate significantly from the majority of observations, may be caused by different mechanisms in relation to normal data [21], [24]–[28]. These different mechanisms include, but are not limited to, factors such as sensor noise, process disturbance and instrument degradation. Relying on a time series tainted by outliers proves to be a source of frustration, as it often leads to misinterpretation and model misspecification. To help offset this aspect of data collection, data pre-processing is a necessary pre-requisite to monitoring data processing [45]. In order to detect and imputing these





**Figure 1.** Original time series with outliers and missing values for Salitre-WWTP (a), GPS (b) and San Fernando-WWTP (c) (self- authorship)

pesky outliers, the Winsorising procedure, which consists of transforming the original data by limiting extreme values, reduces the effect of possibly spurious outliers [46]. Moreover, Winsorising is known for promoting a smoother filter, allowing researchers to place more weight on the central value of each time window by using a mobile window of values ( $N$ ) [47], [48]. Such values depend on  $r$ , the number of values to be modified, where  $2r + 1$  equals window size ( $N$ ); yet, here it is incumbent upon researchers to set the range  $m$  of values to be analysed, for  $m$  is required to establish the minimum and maximum values for the window range and allow the removal of values beyond the upper or lower limits with regard to  $m$  in order to obtain the winsorised data [47]. The process carried out in this investigation can be explained in terms of two multi-part steps.

The first consists of sorting the data in the selected window from lowest to highest; several visual reviews were done to test values of  $r$  and  $m$  parameters. These values were defined once was observed the removal of outliers and before the shape of the signal has lost. Then, values below the minimum acceptable value ( $Min_{AV}$ ) are removed.  $Min_{AV}$  is understood as the value of  $(m + 1)th$  place for all values in the window.  $Min_{AV}$  is represented by Equation 1.

$$Z_i = \begin{cases} X_i & \text{if } X_i > Min_{AV} \\ Min_{AV} & \text{otherwise} \end{cases} \quad (1)$$

where  $X_i$  is the  $i$ th window value,  $Min_{AV}$  the minimum acceptable value and  $Z_i$  the resulting values after the removal of values below  $Min_{AV}$ . As a result, any value lower (outlier) than  $Min_{AV}$  will be replaced. A similar procedure is followed to remove outliers above a threshold  $Max_{AV}$ , which is equal to the value of the  $(N - m)th$  place of all values in the  $Z_i$  window (see Equation 2).

$$Y_i = \begin{cases} Z_i & \text{if } Z_i < Max_{AV} \\ Max_{AV} & \text{otherwise} \end{cases} \quad (2)$$

where  $Z_i$  is the  $i$ th window value,  $Max_{AV}$  the maximum acceptable value and  $Y_i$  the resulting values after removal. It is worth mentioning here that the above, Equation 1 and Equation 2, adapts the method proposed by [46].

Once the corrected time series (without outliers) is obtained, a DFT procedure to complete the time series can be applied. With DFT, the time series is calculated to facilitate the switch from the time domain to the frequency domain. This technique involves the conversion of a finite number of equally spaced samples (discrete points) into a number of coefficients that stem from a finite combination of complex sinusoid components. Doing so ensures that the frequency domain is comprised of the same number of sample values as the previous time domain [48], [49].

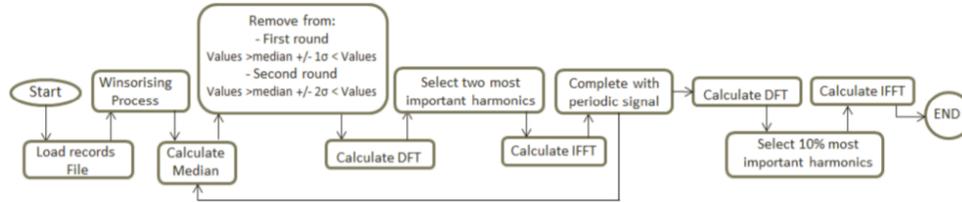
Equation 3 details this conversion. DFT likewise ranks components based on their importance, with importance determined by component amplitude. After ranking, components are eliminated from lower to higher importance, leaving only the most significant harmonics (10 %) in the resulting values. Finally, the Inverse Fast Fourier Transform (IFFT) is utilized to convert complex sinusoids (harmonics) into a finite number of discrete points. This movement represents a return to the time domain (as opposed to the time to frequency shift found in the initial DFT process), as shown in Equation 4. For more details, see [50].

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j2\pi \frac{k}{N}n} \quad k = 0, 1, \dots, N-1 \quad (3)$$

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{-j2\pi \frac{k}{N}n} \quad n = 0, 1, \dots, N-1 \quad (4)$$

where  $n$  is the current sample that is analysed,  $N$  is the total number of time samples taken,  $k$  is current harmonic (frequency) that is considered (0 Hertz up to  $N-1$  Hertz),  $x_n$  is the value of the time series at time  $n$  and  $X_k$  is the amount of frequency  $k$  in the signal (complex number) [49].

The Winsorising process (first step) is applied to the absorbance time series. However, outlier values may still remain; these values must be dealt with. Having outlined the first step, we can proceed to the second step. The procedure obtains the median value from the entire dataset; next, any upper and lower values within median value  $\pm$  one standard deviation range will be discarded (referred to round one). In fact, one standard deviation is used because the time series is highly affected by outlier values that remain after the first round and before DFT is applied. Consequently, the dataset must be switched from the time to the frequency domain. This switch, done with values present in the data set prior to the first missing values gap, naturally leads to the selection of the two most important harmonics. These two harmonics are selected “naturally” insofar as they reproduce the pattern and dynamic of the events. In effect, it is possible to reproduce the outliers if more than two harmonics are included. By virtue of IFFT, the data is translated from the frequency domain back into the time domain. Afterwards, the resulting time series is used to complete the first gap of missing values. Then, the median value from the entire dataset is again obtained and any upper and



**Figure 2.** Flow diagram depicting detection and removal of outliers and replacement of missing values (self-authorship)

lower value within median value +/- two standard deviations range will be taken out (referred to as round two).

On account of Winsorising, along with the two rounds of standard deviation refinement, the effect of outliers at this stage has been attenuated but not fully counteracted. So, the DFT process is run again with only the two most important harmonics applied, followed by another run of IFFT. The resultant time domain values are then able to facilitate the completion of the first missing values gap. When this process is completed, the first gap (missing values) will have been complemented. Therefore, the process described up to this point is brought to bear on the corrected data ranging from the starting point (including the first “completed” gap) to the beginning of the second gap. The same is done with all the corrected data from the initial point to the start of the third gap (including the “completed” first and second gaps) and so on.

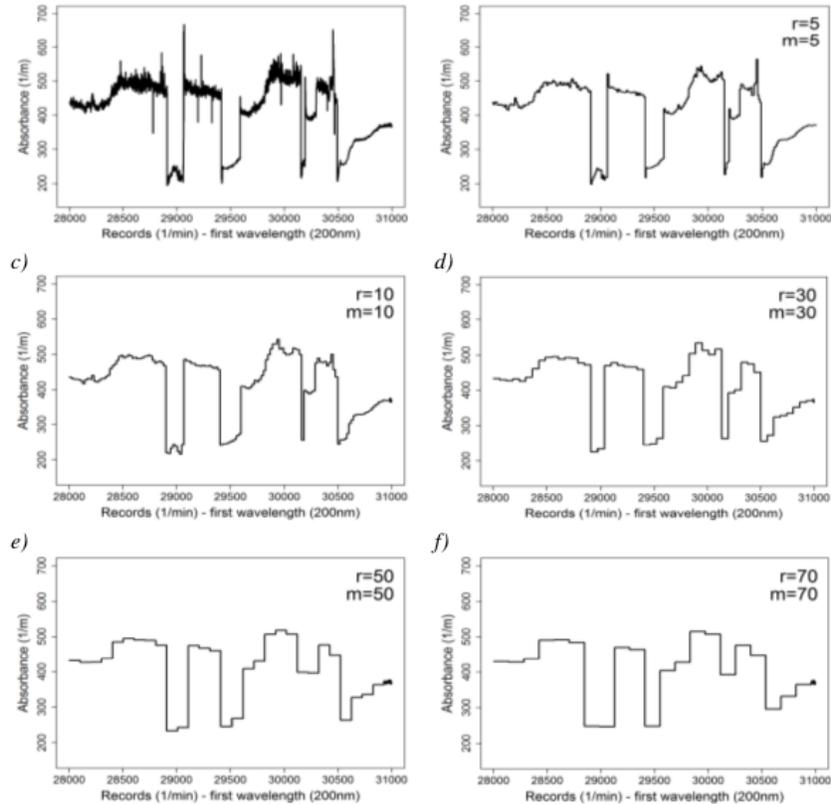
Finally, DFT is applied to the resulting time series and the 10 % of the most important harmonics are used. Going from the frequency domain back to the time domain is once again done with the IFFT process. At this stage, new values imputing either outliers or missing values have the same, or almost the same, shape as the original time series, granting the “macro” vision of the time series coherence. To follow the intricacies of this process, readers can consult the flow diagram shown in Figure 2.

However, as would be expected, the proposed procedure was tested to determine its validity. To this end, continuous subsets of absorbance time series with no outliers and no missing values were intentionally removed from the original time series. In order to assess the accuracy of the method proposed, the mean absolute percentage of error (MAPE) was used as shown in Equation 5, where Val real is the original time series value and Val rep is the replaced time series value.

$$MAPE = \frac{\sum_{i=1}^n \frac{|Val_{i-real} - Val_{i-rep}|}{Val_{i-real}}}{n} \times 100\% \quad (5)$$

### 3. Results and Discussion

For all the study sites, the Winsorising process, coded with *R* statistical language [51], was applied using a value of ten (10) for *r* and *m* parameters to build the window size and find the minimum and maximum acceptable values. Several value combinations of *r* and *m* parameters were verified, and the best results using *r* and *m* proved to be a value of ten (10). Although

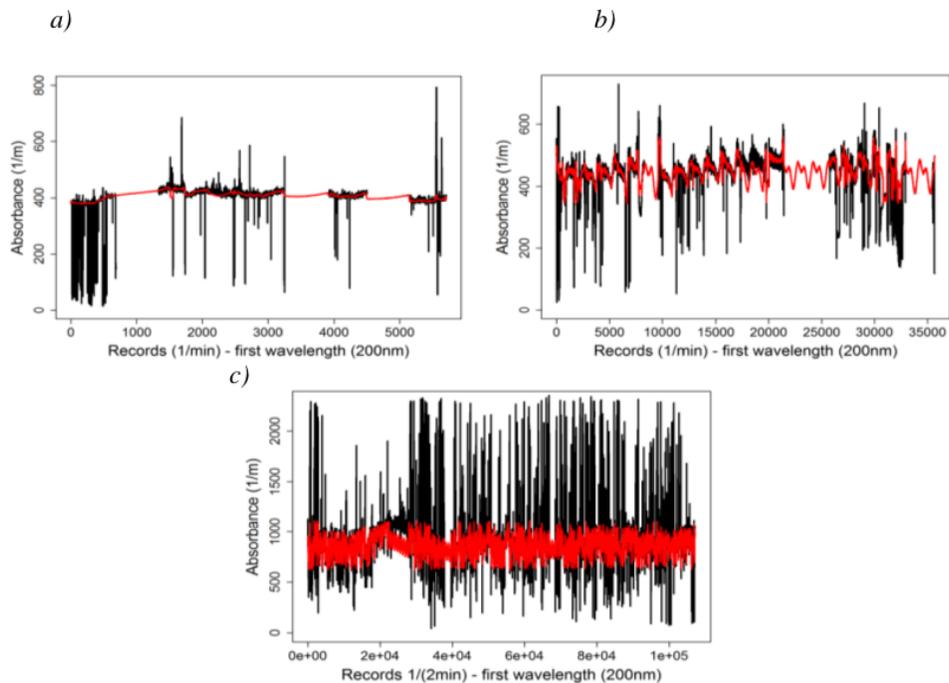


**Figure 3.** Comparison of different  $r$  and  $m$  Winsorising parameter values for the GPS Absorbance time series (section spanning values 28000 to 31000) with original (a),  $r = m = 5$  selected value (b),  $r = m = 10$  selected value (c),  $r = m = 30$  selected value (d) and  $r = m = 50$  selected value (e),  $r = m = 70$  selected value (f) (self-authorship)

using a bigger window size with 20, 30, 50 or 100 as values for the  $m$  parameter, in addition to a lower  $r$  value compared with  $m$ , effectively eliminates outliers, but the shape of the resulting time series is reminiscent of stair steps (thereby losing its resemblance to the original shape of the time series—Figure 3).

Figure 3 shows a select section of the GPS absorbance time series, which includes values from 28000 to 31000 for the time series. Five different combinations for  $r$  and  $m$  parameter values were used, which led to the observation of a time series that mirrors the original time series analysed.

Figure 3 demonstrates that although large window size values (30, 50 or 70) for  $r$  and  $m$  parameters translate into less outlier values, the time series' shape becomes stair-like, losing its similarity to the shape of the original times series (Figure 3, graph a). Thus, using these large window size values would be disadvantageous because an analysis of the last part of the time series (roughly values 30950 until 31000) becomes largely impossible, as observed in Figure 3 (graphs d, e and f).



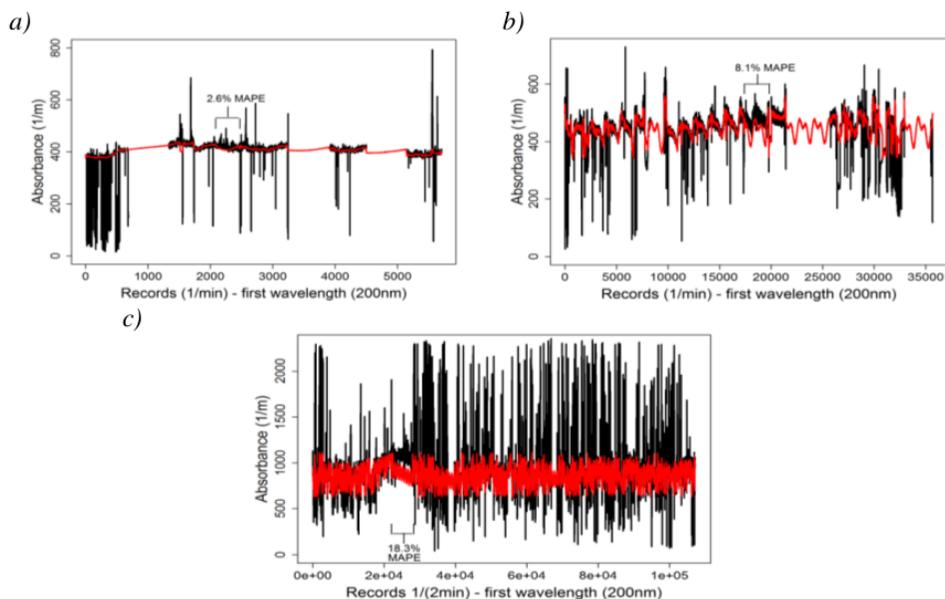
**Figure 4.** Resulting time series for all three study sites after applying the proposed procedure to Salitre-WWTP (a), GPS (b) and San Fernando- WWTP (c) (self-authorship)

However, Figure 3 (graph b) shows the results with some outliers. Therefore, 10 was selected as the optimal value for  $r$  and  $m$  parameters to strike a balance between the analysed window size and the number of outlier values requiring removal. However, given the fact that some outlier values still remain after applying the Winsorising process, the process of applying one and two standard deviations and DFT is repeated to expunge these outliers from the dataset.

As a result, after the final parameter values are selected for window size using the Winsorising process, the last part of the processes indicated in the flow diagram are undertaken (Figure 2). The time series for all three study sites after applying this complete process can be seen in Figure 4.

Figure 4's black curve represents the original times series (absorbance) and its red curve the resulting time series. The results obtained are as follows:

- a) Salitre-WWTP: three large gaps of missing values. Although DFT could not be used for imputing missing values in the first gap, it was successfully used for the other two large gaps. Instead of DFT for the first gap, a linear interpolation was used.
- b) GPS: three large gaps of missing values imputed with DFT.
- c) San Fernando-WWTP: longest absorbance time series and greatest presence of outliers (of the three study sites). The entire filter procedure (Winsorising, one and two standard deviations,



**Figure 5.** Resulting time series and MAPE values for all three study sites after removal a selecting range values and applying the proposed procedure to Salitre-WWTP (a), GPS (b) and San Fernando-WWTP (c) (self-authorship)

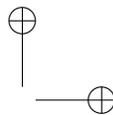
DFT, IFFT) detected and removed outliers, with DFT reapplied for imputing missing values (see Figure 4, graph c).

To assess the validity of the results presented by the proposed methodology, continuous subsets (a section) of the absorbance time series without outlier or missing values were removed from the original time series. Similarly, the MAPE (see Equation 5) was used, as shown in Figure 5. Thus, DFT is applied to the resulting time series and according to preliminary test varying from 5 % to 20 % of the most important harmonics (see the appendix). Therefore, it was decided to use the 10 % of the most important harmonics as a sufficient number of harmonics in order to capture de variability of the time series.

Figure 5 shows the time series (black curve) and the modified-filtered time series (red curve) for all three study sites. Below are the results.

- Salitre-WWTP: values 2100 to 2449 were removed, which gives a 2,6 % MAPE.
- GPS: values 17500 to 19499 were removed, which gives an 8,1 % MAPE.
- San Fernando-WWTP: values 22500 to 27999 were removed, which gives an 18,3 % MAPE.

As the above paragraphs a-c witness, different combinations of values were intentionally eliminated from the original time series (absorbance) at each of the three study sites, in the line with the aforementioned criteria for value exclusion. Overall, a 12 % average of MAPE was obtained.



## 4. Conclusions

The main contribution of this work is the proposal of a new methodology in order to complete time series with missing values and different characteristics (absorbance), and consists of Winsorising as a step in outlier removal. Also, the application of the DFT and the IFFT, using the 10 % most important harmonics of useful values, which is crucial for its later use in different applications, specifically for time series of water quality and quantity in urban sewer systems. The imputing of missing values process (DFT and IFFT) can be applied to any missing value gap with different length. It was applied the combination of the Winsorising and DFT procedures to perform the task of detection, removal of outliers and imputing of missing values to maintain the observed periodic behaviour of these time series.

The process laid out in this paper consists of Winsorising as a first step in outlier removal. One and two standard deviations were then removed (with the concomitant applications of DFT and IFFT) before repeating DFT and IFFT. As a whole, this process effectively removes all outliers and fills the gaps in the time series.

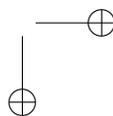
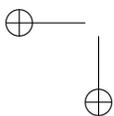
Likewise, the proposed process relies on small  $r$  and  $m$  parameter values (where  $r = m$ ) for window sizes. The three different time series (absorbance) to which the process was applied each exhibited different behaviours and had different sizes. Nevertheless, good results were obtained in terms of what researchers would expect the time series to look like even though there is no dataset with which to compare the time series. Additionally, these results bode well for future applicability in that they offer themselves for application in correcting other hydrologic time series.

DFT allows for the completion of the time series based on the fact that it includes various gap sizes, removes outliers and imputing of missing values. DFT meant lower error percentages at all three study sites, with an error average of 12 %. This reflects what would have likely been the shape or pattern of the time series behaviour had there not been outliers or missing values in the first place.

This work is the starting point to continue with the spectral density estimation process by means of modified rectangular window averaging periodograms with a 50 % overlapping to estimate the periodic behaviour that could be hidden. Also, apply the proposed methodology to different time series with different behaviour and different length as rainfall information, pH, conductivity, temperature, etc. On the other hand, it is necessary to compare the performance of the DFT with other techniques, as machine learning technics (ANN, SVM, AG, etc.) to detect, remove outlier values and impute missing values for time series.

## 5. Acknowledgement

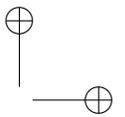
Authors acknowledge Bogotá Water and Sewage Company (Empresa de Acueducto y Alcantarillado de Bogotá – EAAB, under the Administrative Contract No. 9-07-25100-0763-2010) and Medellín Water and Sewage Company (Empresas Públicas de Medellín – EPM) for providing the information used in this research. Also, thank the reviewers and editors for suggestions and observations that have improved the quality of the manuscript.



## Referencias

- [1] Langergraber, G., Fleischmann, N., ofstaedter, F. and Weingartner A., "Monitoring of a paper mill waste water treatment plant using UV/VIS spectroscopy". *IWA Water Science and Technology*, 49(1), 2004, pp. 9-14. ↑112, 113
- [2] Youquan, Z., Yuchun, L., Yang, Z. and Yanjun, F., "A Novel Monitoring System for COD Using Optical Ultraviolet Absorption Method". *Procedia Environmental Sciences*, 10, 2011, pp. 2348-2353. ↑112
- [3] Storey, M., van der Gaag, B. and Burns, B., "Advances in on-line drinking water quality monitoring and early warning systems". *Water Research*, 45, 2011, pp. 741-747. ↑112
- [4] Sempere-Payá, V. and Santonja-Climent, S., "Integrated sensor and management system for urban waste water networks and prevention of critical situations". *Computers, Environment and Urban Systems*, 36, 2012, pp. 65-80. ↑112
- [5] Xu, Z., Liu, B., Dong, Q., Lei, Y., Li, Y., Ren, J., and McCutcheon, J., "Flat microliter membrane-based microbial fuel cell as "on-line sticker sensor" for self-supported in situ monitoring of wastewater shocks". *Bioresource Technology*, 197, 2015, pp. 244-251. ↑112
- [6] Bowerman, B., O'Connell, R. and Koehler, A., *Forecasting, Time Series, and Regression: An Applied Approach*. Fourth Edition. Thomson Learning. USA 2006. ↑112
- [7] Gujarati, D. and Porter, D., *Basic Econometrics*. Fifth Edition. McGraw-Hill Higher Education/Irwin New York-USA. 2008. ↑112
- [8] Lind, D., Marchal, W. and Wathen, S., *Statistical techniques in business & economics*. Fifteenth Edition. McGraw-Hill/Irwin. New York-USA 2012. ↑112
- [9] Drolc, A. and Vrtovšek, J., "Nitrate and nitrite nitrogen determination in waste water using on-line UV spectrometric method". *Bioresource Technology*, 101, 2010, pp. 4228-4233. ↑112
- [10] Al-Monami, F. and Örmeci, B., "Measurement of polyacrylamide polymers in water and wastewater using an in-line UV-vis spectrophotometer". *Journal of Environmental Chemical Engineering*, 2, 2014, pp. 765-772. ↑112
- [11] Bollmann, U., Vollertsen, J., Carmeliet, J. and Bester, K., "Dynamics of biocide emissions from buildings in a suburban stormwater catchment – Concentrations, mass loads and emission processes". *Water Research*, 56, 2014, pp. 66-76. ↑112
- [12] Altmann, J., Massa, L., Sperlich, A. and Gnirss, R., "UV254 absorbance as real-time monitoring and control parameter for micropollutant removal in advanced wastewater treatment with powdered activated carbon". *Water Research*, 94, 2016, pp. 240-245. ↑112
- [13] Murla, D., Gutierrez, O., Martinez, N., Suárez, D., Malgrat, P. and Poch, M., "Coordinated management of combined sewer overflows by means of environmental decision support systems". *Science of the Total Environment*, 550, 2016, pp. 256-264. ↑112
- [14] Lacour, C., Joannis, C. and Chebbo, G., "Assessment of annual pollutant loads in combined sewers from continuous turbidity measurements: Sensitivity to calibration data". *Water Research*, 43, 2009, pp. 2179-2190. ↑112
- [15] Becouze-Lareure, C., Thiebaud, I., Bazin, C., Namour, P., Breil, P. and Perrodin, Y., "Dynamics of toxicity within different compartments of a peri-urban river subject to combined sewer overflow discharges". *Science of the Total Environment*, 539, 2016, pp. 503-514. ↑112
- [16] Gasperi, J., Gromaire, M., Kafi, M., Moilleron, R. and Chebbo, G., "Contributions of wastewater, runoff and sewer deposit erosion to wet weather pollutant loads in combined sewer systems". *Water Research*, 44, 2010, pp. 5875-5886. ↑112
- [17] Métadier, M., & Bertrand-Krajewski, J.-L., "Assessing dry weather flow contribution in TSS and COD storm events loads in combined sewer systems". *Water Science and Technology*, 63(12), 2011, pp. 2983-2991. ↑112
- [18] Bi, E., Monette, F. and Gasperi, J., "Analysis of the influence of rainfall variables on urban effluents concentrations and fluxes in wet weather". *Journal of Hydrology*, 523, 2015, pp. 320-332. ↑112
- [19] Saagi, R., Flores-Alsina, X., Fu, G., Butler, D. and Gernaey, K., "Catchment & sewer network simulation model to benchmark control strategies within urban wastewater systems". *Environmental Modelling & Software*, 78, 2016, pp. 16-30. ↑112
- [20] Johnson, R. and Wichern, D., *Applied Multivariate Statistical Analysis*. 6th ed. Pearson Prentice Hall, USA, 2007. ↑112
- [21] Díaz, C., García, P., Alonso, J., Torres, J. and Taboada, J., "Detection of outliers in water quality monitoring samples using functional data analysis in San Esteban estuary (Northern Spain)". *Science of the Total Environment*, 439, 2012, pp. 54-61. ↑112, 113

- [22] Cucina, D., di Salvatore, A. and Protopapas, M., "Outliers detection in multivariate time series using genetic algorithms". *Chemometrics and Intelligent Laboratory Systems*, 132, 2014, pp. 103-110. ↑112, 113
- [23] Gharibnezhad, F., Mujica, L. and Rodellar, J., "Applying robust variant of Principal Component Analysis as a damage detector in the presence of outliers". *Mechanical Systems and Signal Processing*, 50-51, 2015, pp. 467-479. ↑112, 113
- [24] Grané, A. and Veiga, H. "Wavelet-based detection of outliers in financial time series". *Computational Statistics and Data Analysis*, 54, 2010, pp. 2580-2593. ↑113
- [25] Piñero, J., Martínez, J. García, P., Alonso, J., Díaz, C. and Taboada, J., "Analysis and detection of outliers in water quality parameters from different automated monitoring stations in the Miño river basin (NWSpain)". *Ecological Engineering*, 60, 2013, pp. 60-66. ↑112, 113
- [26] Gumedeze, F. and Chatora, T., "Detection of outliers in longitudinal count data via overdispersion". *Computational Statistics and Data Analysis*, 79, 2014, pp. 192-202. ↑113
- [27] Sangeux, M. and Polak, J., "A simple method to choose the most representative stride and detect outliers". *Gait & Posture*, 41, 2015, pp. 726-730. ↑113
- [28] Qi, M., Fu, Z. and Chen, F., "Outliers detection method of multiple measuring points of parameters in power plant units". *Applied Thermal Engineering*, 85, 2015, pp. 297-303. ↑112, 113
- [29] Martínez, J., Saavedra, A., García-Nieto, P., Piñero, J., Iglesias, C., Taboada, J., Sancho, J. and Pastor, J., "Air quality parameters outliers detection using functional data analysis in the Langreo urban area (Northern Spain)". *Applied Mathematics and Computation*, 241, 2014, pp. 1-10. ↑113
- [30] Maciá-pérez, F., Berna-Martínez, J., Fernández, A. and Abreu, M., "Algorithm for the detection of outliers based on the theory of rough sets". *Decision Support Systems*, 75, 2015, pp. 63-75. ↑113
- [31] Song, X., Liu, Z., Yang, J. and Qi, Y., "Extended semi-supervised fuzzy learning method for nonlinear outliers via pattern discovery". *Applied Soft Computing*, 29, 2015, pp. 245-255. ↑113
- [32] Dumedah, G., Jeffrey, P. and Li, W., "Assessing artificial neural networks and statistical methods for infilling missing soil moisture records". *Journal of Hydrology*, 515, 2014, pp. 330-344. ↑113
- [33] Figueroa, J., Kalenatic, D. and Lopez, C., "Missing data imputation in multivariate data by evolutionary algorithms". *Computers in Human Behavior*, 27(5), 2011, pp. 1468-1474. ↑113
- [34] Figueroa-García, J., Kalenatic, D. and Lopez, C., "Incomplete Time Series: Imputation through Genetic Algorithms". *Time Series Analysis, Modeling and Applications*, 47, 2013, pp. 31-52. ↑113
- [35] de França, F., Coelho, G. and von Zuben, F., "Predicting missing values with biclustering: A coherence-based approach". *Pattern Recognition*, 46, 2013, pp. 1255-1266. ↑113
- [36] Folch-Fortuny, A., Arteaga, F. and Ferrer, A., "PCA model building with missing data: New proposals and a comparative study". *Chemometrics and Intelligent Laboratory Systems*, 146, 2015, pp. 77-88. ↑113
- [37] Carvajal, C., Bayona, D. and Ortiz, Z., "Taxonomy extension and missing-values treatment over an informatics-security incident repository". *Ingeniería*, 18(1), 2013, pp. 24-49. ↑113
- [38] Dumedah, G. and Coulibaly, P., "Evaluation of statistical methods for infilling missing values in high-resolution soil moisture data". *Journal of Hydrology*, 400, 2011, pp. 95-102. ↑113
- [39] Haworth, J. and Cheng, T., "Non-parametric regression for space-time forecasting under missing data". *Computers, Environment and Urban Systems*, 36, 2012, pp. 538-550. ↑113
- [40] Junger, W. and Ponce, A., "Imputation of missing data in time series for air pollutants". *Atmospheric Environment*, 102, 2015, pp. 96-104. ↑113
- [41] Avellaneda, J., Ochoa, C., and Figueroa-García, J., "Comparison between a self organizing neural fuzzy system and an ARIMAX model to forecasting volatile economic series". *Ingeniería*, 17(2), 2012, pp. 26-34. ↑113
- [42] Tukey, J., *Exploratory data analysis*. Addison-Wesely. ↑113
- [43] Acuña, E. and Rodríguez, C., *On Detection of Outliers and Their Effect in Supervised Classification*. Department of Mathematics University of Puerto Rico at Mayaguez, Mayaguez, Puerto Rico. 2013. [Online]. Available <http://academic.uprm.edu/eacuna/vene31.pdf> ↑113
- [44] s:can, *Manual ana::pro Version 5.3 September 2006 Release*. Messtechnik GmbH, Vienna, Austria, 2006. ↑113
- [45] Liu, H., Shah, S. and Jiang, W., "On-line outlier detection and data cleaning". *Computers and Chemical Engineering*, 28, 2004, pp. 1635-1647. ↑113
- [46] Ko, S-J and Lee, Y., "Theoretical analysis of winsorizing smoothers and their applications to image processing". *Acoustics, Speech, and Signal Processing, ICASSP-1991*, pp. 3001-3004. ↑114, 115
- [47] Pearson, R., "Outliers in process modelling and identification". *IEEE Transactions on Control Systems Technology*, 10, 2002, pp. 55-63. ↑114



- [48] Kontaki, M., Gounaris, A., Papadopoulos, A., Tsihlias, K. and Manolopoulos, Y., "Efficient and flexible algorithms for monitoring distance based outliers over data streams". *Information Systems*, 55, 2015, pp. 37–53. ↑ 114, 115
- [49] Proakis, J. and Manolakis, D., *Digital signal processing principles, algorithms, and applications*. Fourth Ed. New Jersey: Pearson Prentice Hall, 2007. ↑ 115
- [50] Plazas-Nossa, L. and Torres, A., "Fourier analysis as a forecasting tool for absorbance time series received by UV-Vis probes installed on urban sewer systems". *Proceedings of 8th International Conference Novatech*, 2013, Lyon, France, 23-27 June 2013. ↑ 115
- [51] R Core Team, "R: A language and environment for statistical computing". *R Foundation for Statistical Computing*, Vienna, Austria, 2014. [Online]. Available URL <http://www.R-project.org/> ↑ 116

---

### Leonardo Plazas Nossa

Ingeniero Electrónico, Magister en Teleinformática, Doctorado en Ingeniería, Universidad Distrital Francisco José de Caldas, Facultad de Ingeniería, Carrera 7 No. 40-53, 3239300 Ext. 2405. Bogotá, Colombia.  
Contacto: [lplazasn@udistrital.edu.co](mailto:lplazasn@udistrital.edu.co)

---

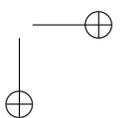
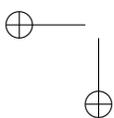
### Miguel Antonio Ávila Angulo

Ingeniero Catastral y Geodesta, Magister en Teleinformática, Universidad Distrital Francisco José de Caldas, Facultad de Ingeniería, Carrera 7 No. 40-53, 3239300 Ext. 2405. Bogotá, Colombia.  
Contacto: [maavila@udistrital.edu.co](mailto:maavila@udistrital.edu.co)

---

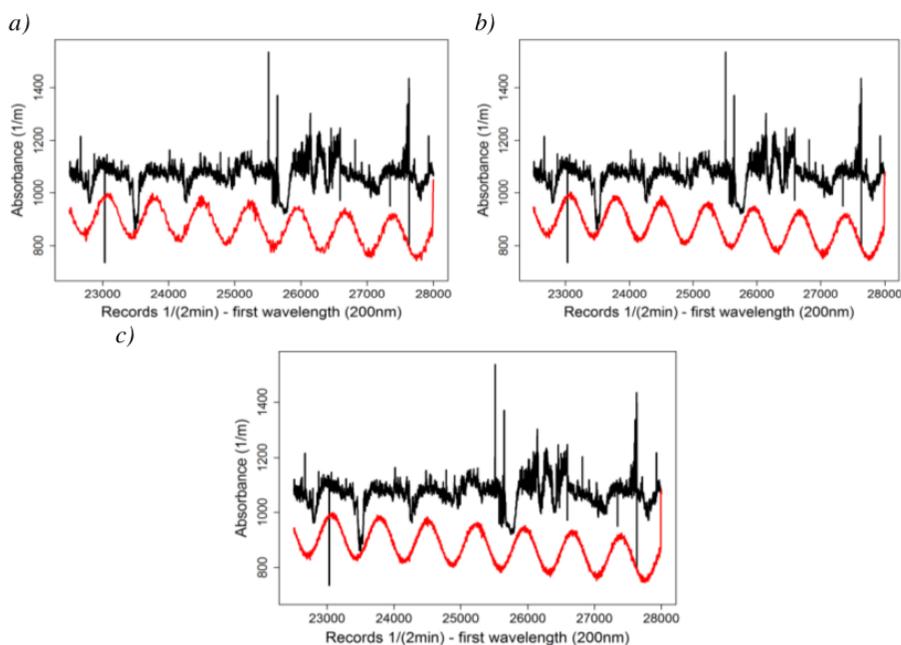
### Andrés Torres

Ingeniero Civil, Especialización en Sistemas Gerenciales de Ingeniería, Maestría en Ingeniería Civil, Doctorado en Ingeniería Civil, Grupo de Investigación Ciencia e Ingeniería del Agua y el Ambiente, Facultad de Ingeniería, Pontificia Universidad Javeriana, Carrera 7 No. 40 -62, 3208320 Ext. 5553. Bogotá, Colombia.  
Contacto: [andres.torres@javeriana.edu.co](mailto:andres.torres@javeriana.edu.co)



## Appendix

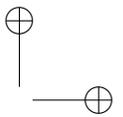
The DFT is applied to the resulting time series, several test were done varying from 5 % to 20 % of the most important harmonics. As an example, Figure 6 shows the results obtained for San Fernando-WWTP applied to the values between 22500 and 27999, which were removed the same process was done for Salitre-WWTP and GPS. Table I shows the MAPE values for 5 %, 10 % and 20 % of the most important harmonics results respectively for all three study sites. Therefore, it was decided to use the 10 % of the most important harmonics as a sufficient number of harmonics in order to capture de variability of the time series.



**Figure 6.** Resulting time series and MAPE values for San Fernando WWTP study site after removal a selecting range values and applying the proposed procedure for 5 % (a), 10 % (b) and 20 % (c) of most relevant harmonics respectively (self-authorship)

**Table I.** MAPE values for 5 %, 10 % and 20 % of the most important harmonics results respectively for all three study sites

Study	MAPE Value		
	5 %	10 %	20 %
Salitre-WWTP	2.59044	2.59056	2.58611
GPS	8.06975	8.07084	8.07012
San Fernando-WWTP	18.22815	18.23058	18.23014



Investigación

## Del Operador Apertura en la Matemática Morfológica Difusa

*On the Fuzzy Opening Morphological Operator*

**Carlos Orlando Ochoa Castillo<sup>1</sup>, Wilson Javier Forero Baquero<sup>2</sup>**

<sup>1</sup>Universidad Distrital Francisco José de Caldas. Bogotá - Colombia, <sup>2</sup>Universidad Nacional de Colombia.

Recibido: 02/08/2016 Modificado: 17/11/2016. Aceptado: 27/12/2016

### Resumen

**Contexto:** Las propiedades que posee un operador de interior se pueden trasladar a la determinación de características específicas de imágenes, las cuales, desde la matemática morfológica difusa pueden llegar a ser analizadas por medio del operador erosión y apertura, al conjugar estas ideas es pertinente indagar en torno a la naturaleza de estos operadores.

**Método:** Gracias a las propiedades reticulares que posee el análisis matemático de una imagen por medio de matemática morfológica difusa, se busca dotar de restricciones al elemento estructural con el cual se desea filtrar la imagen para así obtener caracterizaciones del operador apertura y erosión.

**Resultados:** Se demuestre que si la relación estructural es reflexiva o \*-antitransitiva el operador apertura es interior; en caso que la relación cumpla ambas, la erosión es un operador interior.

**Conclusiones:** Los operadores morfológicos difusos permiten obtener información relevante sin alterar la estructural global de la imagen; arrojando mayor calidad que los métodos clásicos, en especial si se emplea el operador apertura difuso con una relación estructural adecuada.

**Palabras clave:** Apertura, erosión, interior, matemática morfológica, operador.

**Idioma:** Español

### Abstract

**Context:** Interior operators have interesting properties that can be used in detecting relevant features in digital images. In this respect, it is pertinent to study the behaviour of the opening and erosion operators from the perspective of fuzzy morphological mathematics.

**Method:** Bearing in mind the reticular properties inherent to the mathematical analysis of an image by fuzzy morphological mathematics we seek to find restrictions on the structural element intended to filter the image so as to obtain characterizations of the opening and erosion operators.

**Results:** We prove that if the structural relationship is reflexive or \*-antitransitive the opening operator is interior. On the other hand, we found that if the relationship meets both erosion is a interior operator.

**Conclusions:** The fuzzy morphological operators can give us more information than the classical methods when we filter an image, especially if we implement the opening operator with a right structural element.

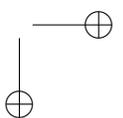
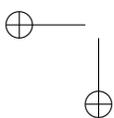
**Keywords:** Erosion, interior, mathematical morphological, opening, operator.

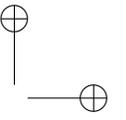
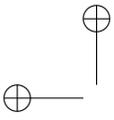
### Open access



Citación: W. J. Forero, C. O. Ochoa, "Del operador apertura en la matemática morfológica difusa" INGENIERÍA, vol. 22, no. 1, pp. 125-139, 2017.

©Los autores; titular de derechos de reproducción Universidad Distrital Francisco José de Caldas. En línea DOI: <http://dx.doi.org/10.14483/udistrital.jour.reveng.2017.1.a08>





## 1. Introducción

El estudio de imágenes diagnósticas por medio de contraste consiste en inyectar una tinta en el sistema circulatorio del paciente observando su flujo y posibles obstrucciones que determinan anomalías en el sujeto; de esta forma se buscan diseñar instrumentos y estrategias para la solución de tal problemática. En la matemática morfológica se han encontrado algunas respuestas de los análisis de las imágenes de operadores, filtros, entre otros [8], [1] pero la lógica clásica implementada en estos procesos limita la calidad de los resultados, como lo estudiado en [2], donde los autores describen problemas de ruido, detección de regiones apropiadas para realizar el análisis, debido a que la imagen está en una escala de grises y no solo en blanco o negro. Es por ello que la lógica difusa se convierte en un instrumento que ofrece la mayor información en la escala que se construya y al aplicar ciertos operadores, se obtendrá la información relevante para el diagnóstico perdiendo la menor cantidad de datos pero eliminando los posibles ruidos.

Un posible camino es estudiar los operadores de erosión y apertura para obtener la mayor cantidad de información en la menor cantidad de iteraciones en el proceso como se sugiere en [8].

## 2. Preliminares

Los operadores de erosión y apertura junto con sus propiedades demandan la presentación de algunos conceptos inherentes a la teoría de retículos y la lógica difusa, entre otros.

### 2.1. Retículos residuados

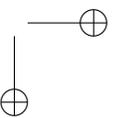
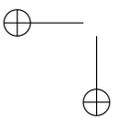
**Definición 1.** Sea  $(L, \leq, \wedge, \vee)$  un retículo y  $*$  una operación binaria definida en  $L$ , se dice que  $(L, \leq, \wedge, \vee, *)$  es un retículo residuado si existe una operación binaria  $\rightsquigarrow$  en  $L$ , que satisface:

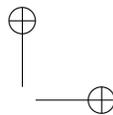
$$\alpha * \beta \leq \gamma \iff \alpha \leq \beta \rightsquigarrow \gamma.$$

En caso de que  $(L, \leq, \wedge, \vee)$  sea un retículo completo, entonces  $(L, \leq, \wedge, \vee, *, \rightsquigarrow, \perp, \top)$  es un retículo residuado completo. La estructura anterior posee las siguientes propiedades

**Lema 1.** [11], [14] Sea  $(L, \leq, \wedge, \vee, *, \rightsquigarrow, \perp, \top)$  un retículo residuado completo, entonces para  $x, y, z$  elementos de  $L$ , se tiene que:

- $x * (x \rightsquigarrow y) \leq y$ ,
- $x \rightsquigarrow (y \rightsquigarrow z) = (x * y) \rightsquigarrow z$ ,
- $x \leq y \iff x \rightsquigarrow y = \top$ ,
- $x * (y \rightsquigarrow z) \leq y \rightsquigarrow (x * z)$ ,
- $x \rightsquigarrow \bigwedge_{i \in I} y_i = \bigwedge_{i \in I} (x \rightsquigarrow y_i)$ ,
- $\bigvee_{i \in I} x_i \rightsquigarrow y = \bigwedge_{i \in I} (x_i \rightsquigarrow y)$ ,
- $x * \bigwedge_{i \in I} y_i = \bigwedge_{i \in I} (x * y_i)$ .





## 2.2. T-normas

Las normas triangulares [15], mejor conocidas como  $t$ -normas permiten una generalización de la intersección entre conjuntos difusos.

**Definición 2.** Una  $t$ -norma es una operación  $*$  :  $[0, 1] \times [0, 1] \rightarrow [0, 1]$  que satisface:

1. Para todo  $x, y, z \in [0, 1]$

$$\begin{aligned}x * y &= y * x \\(x * y) * z &= x * (y * z)\end{aligned}$$

2.  $*$  no es decreciente en ambos argumentos, es decir,

$$\begin{aligned}x_1 \leq x_2 \text{ implica que } x_1 * y &\leq x_2 * y \\y_1 \leq y_2 \text{ implica que } x * y_1 &\leq x * y_2\end{aligned}$$

3.  $1 * x = x$  y  $0 * x = 0$  para todo  $x$  en  $[0, 1]$

El concepto anterior, se relaciona con el de continuidad en  $[0, 1] \times [0, 1]$ , tal como se presenta en el siguiente,

**Lema 2.** [3] Sea  $f : [0, 1] \times [0, 1] \rightarrow [0, 1]$  una operación binaria, si  $f$  es no decreciente en  $x$  entonces  $f$  es continua a izquierda en  $x$  si y solo si para cualquier  $\{a_j | j \in J\}$  y  $b \in [0, 1]$  se cumple que:

$$f\left(\bigvee_{j \in J} a_j, b\right) = \bigvee_{j \in J} f(a_j, b)$$

De la misma forma,  $f$  es continua a derecha si:

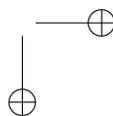
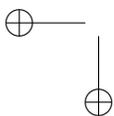
$$f\left(\bigwedge_{j \in J} a_j, b\right) = \bigwedge_{j \in J} f(a_j, b).$$

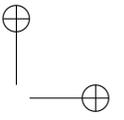
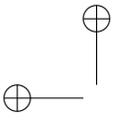
Dada la estructura de orden y topológica que el intervalo  $[0, 1]$  posee como subespacio de  $\mathbb{R}$ , es posible considerar la naturaleza de  $*$  con respecto al concepto de continuidad; desde esta perspectiva, siguiendo a [3] se tiene que  $([0, 1], \leq, \wedge, \vee, *, \rightsquigarrow, 0, 1)$  es un retículo residuado completo con  $*$  una  $t$ -norma continua a izquierda.

## 2.3. Relaciones difusas

En el desarrollo de la matemática difusa uno de los conceptos fundamentales es la de membresía o pertenencia de un elemento a un conjunto, el cual no está determinado únicamente con 0, 1 sino que es multivaluado [11] modificando así la concepción de algunos conceptos entre los cuales se encuentran las relaciones internas de un conjunto.

Se define una relación difusa  $R$  mediante una función  $R : X \times X \rightarrow L$ ; al igual que en el caso clásico es posible caracterizar las propiedades que posea, es por ello que se extiende el concepto de preorden al ambiente difuso, con  $(L, \leq, \wedge, \vee, \top, \perp, *)$  un retículo residuado completo (se entenderá por  $\top$  y  $\perp$ , el elemento máximo y mínimo respectivamente de la estructura).





**Definición 3.** Sea  $(L, \leq, \wedge, \vee, \top, \perp, *)$  y  $X$  un conjunto, una relación difusa (binaria)  $R$  en  $X$  es un  $*$ -preorden difuso si satisface:

1.  $R(x, x) = \top$  para todo  $x \in X$  (Reflexiva),
2.  $R(x, y) * R(y, z) \leq R(x, z)$  para todo  $x, y, z \in X$  (Transitiva).

Además del concepto de  $*$ -preorden difuso, aparecen de manera natural los conceptos de relación  $*$ -tolerante cuando  $R$  es reflexiva y simétrica i.e.  $R(x, y) = R(y, x)$  para todo  $x, y \in X$ ;  $R$  es una relación  $*$ -equivalente si es un  $*$ -preorden difuso y es simétrica, junto a ello  $R^{op}$  es la relación opuesta de  $R$ , i.e.  $R^{op}(x, y) = R(y, x)$ , en caso que se satisfaga que  $R(x, y) * R(y, z) \geq R(x, z)$  se dirá que  $R$  es antitransitiva respecto a  $*$ .

### 3. Algunos operadores morfológicos difusos

Por medio de los operadores morfológicos es posible abarcar el reconocimiento de imágenes con el estudio de ciertos detalles específicos sin llegar a alterar la imagen global, sino actuando en elementos estructurales que se definan previamente; esto se logra a partir de los operadores de dilatación y erosión, los cuales son los elementos constitutivos de cualquier otro operador que se pueda definir. En las líneas que siguen, se entiende por *elemento estructural* a un subconjunto del conjunto en estudio.

En [9] se definen los operadores erosión y dilatación en  $X = \mathbb{R}^2$  o  $X = \mathbb{Z}^2$  respectivamente por,

$$\varepsilon_B(A) = \{y \in X \mid \bar{B}_y \subset A\} \quad (1)$$

$$\delta_B(A) = \{x \in X \mid B_x \cap A \neq \emptyset\} \quad (2)$$

donde  $\bar{B}_y = \{y - b \mid b \in B\}$  y  $B_x = \{x + b \mid b \in B\}$ .

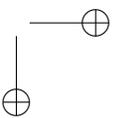
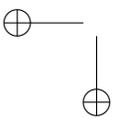
Se extienden las ideas anteriores tomando un conjunto arbitrario  $X$  y una relación difusa  $R$  la cual es la proyección del concepto de elemento estructural y en consecuencia se denomina relación estructural, así las ecuaciones (1) y (2) se definen en un contexto difuso por medio de  $*$  y  $\rightsquigarrow$ , se denota a  $L^X$  como el conjunto de todas las funciones de  $X$  en  $L$ , esto es, el conjunto de todos los subconjuntos  $L$ -difusos de  $X$ .

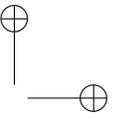
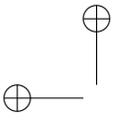
**Definición 4.** Dada una relación difusa  $R \in L^{X \times X}$ , los operadores erosión y dilatación de  $\mu \in L^X$  son,

$$\varepsilon_R(\mu)(x) = (R^{op} \triangleleft \mu)(x) = \bigwedge_{y \in X} \{R(y, x) \rightsquigarrow \mu(y)\}$$

$$\delta_R(\mu)(x) = (R \circ \mu)(x) = \bigvee_{y \in X} \{R(x, y) * \mu(y)\}.$$

Con la composición de los operadores antes definidos es posible crear operadores que nos brinden mayor exactitud en el reconocimiento de detalles específicos de una imagen [9], uno de ellos es el *operador apertura* el cual consta de la composición del operador erosión seguido del operador dilatación, i.e.  $\alpha_R(\mu) = R \circ (R^{op} \triangleleft \mu)$ ; dicho operador es usado especialmente para eliminar regiones





pequeñas y protuberancias.

Si se sustituye a  $\mu \in L^X$  por  $S \in L^{X \times X}$  los conceptos previos se extienden como sigue,

**Definición 5.** Sean  $(L, \leq, \wedge, \vee, \top, \perp, *)$ ,  $X$  un conjunto y  $R, S$  relaciones difusas en  $X$ , Las operaciones  $\circ^*$  y  $\triangleleft_{\rightsquigarrow_*}$  entre  $R$  y  $S$  se definen como,

$$(R \circ^* S)(x, y) = \bigvee_{w \in X} \{R(x, w) * S(w, y)\}$$
$$(R \triangleleft_{\rightsquigarrow_*} S)(x, y) = \bigwedge_{w \in X} \{R(x, w) \rightsquigarrow_* S(w, y)\}$$

Por lo demostrado en [9] se tiene que  $\circ^*$  y  $\triangleleft_{\rightsquigarrow_*}$  forman un par adjunto, esto es:

**Teorema 1.** ([9]) Sean  $(L, \leq, \wedge, \vee, \top, \perp, *)$ ,  $X$  un conjunto y  $Q, R, S$  relaciones binarias difusas en  $X$ , entonces

$$Q \circ^* R \leq S \Leftrightarrow R \leq Q^{op} \triangleleft_{\rightsquigarrow_*} S.$$

## 4. Relación estructural en los operadores de interior morfológicos difusos

Al abordar el estudio de las propiedades de los operadores apertura y erosión definidos en la sección anterior, se presentan otros objetos relacionados.

**Definición 6.** ([13], [5]) Un operador  $C : L^X \rightarrow L^X$  es un operador interior difuso en  $X$  si se verifica:

1.  $C(\mu) \leq \mu$  para todo subconjunto difuso  $\mu \in L^X$ ,
2.  $\mu_1 \leq \mu_2 \Rightarrow C(\mu_1) \leq C(\mu_2)$  para todo  $\mu_1, \mu_2 \in L^X$ ,
3.  $C(C(\mu)) = C(\mu)$  para todo  $\mu \in L^X$ .

Se presenta el concepto de operador  $*$ -coherente con  $\hat{a}(x) = 1$  si  $x = a$  y  $\hat{a}(x) = 0$  en otro caso.

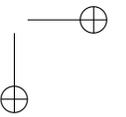
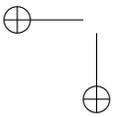
**Definición 7.** Sean  $(L, \leq, \wedge, \vee, \top, \perp, *)$ ,  $X$  un conjunto y un operador  $C : L^X \rightarrow L^X$ ; se dice que  $C$  es  $*$ -coherente si

$$\mu(a) * C(\hat{a})(x) \leq C(\mu)(x)$$

para todo  $\mu \in L^X$  y para todo  $(a, x) \in X \times X$ .

Lo anterior conlleva el siguiente enunciado,

**Teorema 2.** Sean  $(L, \leq, \wedge, \vee, \top, \perp, *)$ ,  $X$  un conjunto y  $R \in L^{X \times X}$ , los operadores  $\alpha_R$  apertura y  $\epsilon_R$  erosión son  $*$ -coherentes.



*Demostración.* Sea  $\mu \in L^X$  y  $x, a \in X$ ,

$$\begin{aligned}
 \mu(a) * \alpha(\hat{a})(x) &= \mu(a) * \bigvee_{b \in X} \{R(x, b) * (R^{op} \triangleleft \hat{a})(b)\} \\
 &= \bigvee_{b \in X} \{\mu(a) * R(x, b) * (R^{op} \triangleleft \hat{a})(b)\} \\
 &= \bigvee_{b \in X} \{R(x, b) * \mu(a) * \bigwedge_{c \in X} \{R^{op}(b, c) \rightsquigarrow \hat{a}(c)\}\} \\
 &\leq \bigvee_{b \in X} \{R(x, b) * \bigwedge_{c \in X} \{\mu(a) * (R^{op}(b, c) \rightsquigarrow \hat{a}(c))\}\} \\
 &\leq \bigvee_{b \in X} \{R(x, b) * \bigwedge_{c \in X} \{R^{op}(b, c) \rightsquigarrow (\mu(a) * \hat{a}(c))\}\} \\
 &\leq \bigvee_{b \in X} \{R(x, b) * \bigwedge_{c \in X} \{R^{op}(b, c) \rightsquigarrow \mu(c)\}\} \\
 &= \bigvee_{b \in X} \{R(x, b) * (R^{op} \triangleleft \mu)(b)\} \\
 &= \alpha(\mu)(x)
 \end{aligned}$$

por lo tanto,

$$\mu(a) * \alpha(\hat{a})(x) \leq \alpha(\mu)(x)$$

De manera análoga se tiene para  $\varepsilon_R$ . □

Es oportuno notar que lo afirmado es independiente de la relación estructural  $R$  que se tome, lo cual no ocurre al estudiar el concepto de interior con los operadores morfológicos en general, es necesario restringir a  $R$ , con lo cual se tiene que,

**Teorema 3.** Si  $R$  es reflexiva y  $R^{op}$  es  $*$ -antitransitiva entonces  $\varepsilon_R$  es un operador interior.

*Demostración.* Si  $\mu \in L^X$  y  $x, a \in L$ ,

■

$$\begin{aligned}
 \varepsilon_R(\mu) &= \bigwedge_{y \in L} \{R^{op}(x, y) \rightsquigarrow \mu(y)\} \\
 &\leq R(x, x) \rightsquigarrow \mu(x) \\
 &= \top \rightsquigarrow \mu(x) \\
 &= \mu(x)
 \end{aligned}$$

es decir  $\varepsilon_R(\mu) \leq \mu$ .

- Debido a que  $\rightsquigarrow$  es monótona en el segundo argumento entonces si  $\mu_1 \leq \mu_2$  implica que  $\varepsilon_R(\mu_1) \leq \varepsilon_R(\mu_2)$ .
- Debido a que las dos condiciones anteriores se satisfacen, solo resta probar la desigualdad  $\varepsilon_R(\mu) \leq \varepsilon_R(\varepsilon_R(\mu))$ ,

$$\begin{aligned}
 \varepsilon_R(\varepsilon_R(\mu)) &= \bigwedge_{y \in L} \{R^{op}(x, y) \rightsquigarrow \varepsilon_R(\mu)(y)\} \\
 &= \bigwedge_{y \in L} \{R^{op}(x, y) \rightsquigarrow \bigwedge_{w \in L} \{R^{op}(y, w) \rightsquigarrow \mu(w)\}\} \\
 &= \bigwedge_{y \in L} \bigwedge_{w \in L} \{R^{op}(x, y) \rightsquigarrow (R^{op}(y, w) \rightsquigarrow \mu(w))\} \\
 &= \bigwedge_{y \in L} \bigwedge_{w \in L} \{(R^{op}(x, y) * R^{op}(y, w)) \rightsquigarrow \mu(w)\} \\
 &\geq \bigwedge_{y \in L} \bigwedge_{w \in L} \{R^{op}(x, w) \rightsquigarrow \mu(w)\} \\
 &= \varepsilon_R(\mu)(x).
 \end{aligned}$$

□

La cuestión natural es ver si las mismas condiciones de  $R$  son necesarias para afirmar que el operador apertura es interior, lo cual tiene por respuesta el siguiente enunciado,

**Teorema 4.** Si  $R$  es reflexiva o  $*$ -antitransitiva entonces  $\alpha_R$  es un operador interior.

*Demostración.* Si  $\mu \in L^X$  y  $x, a \in L$ ,

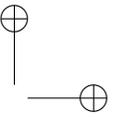
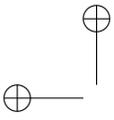
▪

$$\begin{aligned}
 \alpha_R(\mu)(x) &= \bigvee_{y \in L} \{R(x, y) * (R^{op} \triangleleft \mu)(y)\} \\
 &\leq \bigvee_{y \in L} \{R(x, y) * (R^{op}(y, x) \rightsquigarrow \mu(x))\} \\
 &\leq \bigvee_{y \in L} \mu(x) \\
 &= \mu(x).
 \end{aligned}$$

por tanto,  $\alpha_R(\mu) \leq \mu$ .

- De la definición de t-norma y la monotonía en el segundo argumento de  $\rightsquigarrow$  se satisface: si  $\mu_1 \leq \mu_2$  entonces  $\alpha_R(\mu_1) \leq \alpha_R(\mu_2)$ .
- Por lo anterior solo resta verificar que  $\alpha_R(\mu) \leq \alpha_R(\alpha_R(\mu))$ , si  $R$  es reflexiva,

$$\begin{aligned}
 \alpha_R(\alpha_R(\mu))(x) &= \bigvee_{y \in L} \{R(x, y) * (R^{op} \triangleleft \mu)(y)\} \\
 &\geq R(x, x) * \alpha_R(\mu)(x) \\
 &= \top * \alpha_R(\mu)(x) \\
 &= \alpha_R(\mu)(x)
 \end{aligned}$$



ahora si  $R$  es  $*$ -antitransitiva,

$$\begin{aligned}\alpha_R(\alpha_R(\mu))(x) &= \bigvee_{y \in L} \{R(x, y) * (R^{op} \triangleleft \mu)(y)\} \\ &\geq R(x, x) * \alpha_R(\mu(x)) \\ &= \bigvee_{y \in L} \{R(x, x) * (R(x, y) * (R^{op} \triangleleft \mu)(y))\} \\ &\geq \bigvee_{y \in L} \{R(x, y) * (R^{op} \triangleleft \mu)(y)\} \\ &= \alpha_R(\mu(x))\end{aligned}$$

por lo tanto, en caso que  $R$  sea reflexiva o  $*$ -antitransitiva se satisface que:

$$\alpha_R(\mu) = \alpha_R(\alpha_R(\mu)).$$

□

Los anteriores resultados, llevan a formularse la cuestión en torno a la relación entre los operadores erosión y apertura, con facilidad se tiene que

**Corolario 1.** Sea  $R \in L^{X \times X}$  se tiene que para todo  $\mu \in L^X$ ,

$$\varepsilon_R(\mu) \leq \alpha_R(\mu).$$

Al profundizar en las relaciones de  $*$ -equivalencia, se encuentra en [6] que el operador erosión posee las siguientes propiedades,

**Lema 3.** Sea  $R$  una relación de  $*$ -equivalencia, se satisfacen las siguientes propiedades,

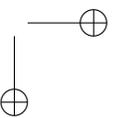
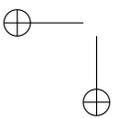
1.  $\varepsilon_R$  es un operador interior,
2.  $\varepsilon_R(\bigwedge_{i \in I} \mu_i) = \bigwedge_{i \in I} (\varepsilon_R(\mu_i))$  para cualquier  $\mu_i \in L^X$ ,
3.  $\varepsilon_R(\gamma \rightsquigarrow \mu) = \gamma \rightsquigarrow \varepsilon_R(\mu)$  para cualquier constante  $\gamma \in [0, 1]$  y  $\mu \in [0, 1]^X$ .

Tratando de formular lo correspondiente al operador apertura, se tiene el siguiente lema,

**Lema 4.** Sea  $R$  una relación de  $*$ -equivalencia, se satisfacen las siguientes propiedades,

1.  $\gamma * \alpha_R(\mu) \leq \alpha_R(\gamma * \mu)$  para cualquier constante  $\gamma \in [0, 1]$  y  $\mu \in [0, 1]^X$ .
2.  $\alpha_R(\gamma \rightsquigarrow \mu) \leq \gamma \rightsquigarrow \alpha_R(\mu)$  para cualquier constante  $\gamma \in [0, 1]$  y  $\mu \in [0, 1]^X$ .
3.  $\alpha_R(\bigwedge_{i \in I} \mu_i) \leq \bigwedge_{i \in I} (\alpha_R(\mu_i))$  para cualquier  $\mu_i \in L^X$ ,

*Demostración.* Con base en las propiedades del corolario 1 y el lema 3,



- Sea  $\mu \in L^X$ ,

$$\begin{aligned}\gamma * \alpha_R(\mu)(x) &= \bigvee_{a \in X} \gamma * (R(x, a) * \bigwedge_{b \in X} R^{op}(a, b) \rightsquigarrow \mu(b)) \\ &\leq \bigvee_{a \in X} R(x, a) * \bigwedge_{b \in X} R^{op}(a, b) \rightsquigarrow (\gamma * \mu(b)) \\ &= \alpha_R(\gamma * \mu).\end{aligned}$$

- Sea  $\mu \in L^X$ ,

$$\begin{aligned}\alpha(\gamma \rightsquigarrow \mu) &= \bigvee_{a \in X} R(x, a) * [R^{op} \triangleleft (\gamma \rightsquigarrow \mu)](a) \\ &= \bigvee_{a \in X} R(x, a) * \gamma \rightsquigarrow R^{op} \triangleleft \mu \\ &\leq \gamma \rightsquigarrow \bigvee_{a \in X} R(x, a) * (R^{op} \triangleleft \mu)(a) \\ &= \gamma \rightsquigarrow \alpha_R(\mu)(x).\end{aligned}$$

- Sea  $\mu \in L^X$ ,

$$\begin{aligned}\alpha_R(\bigwedge_{i \in I} \mu_i(x)) &= \bigwedge_{a \in X} R(x, a) * (R^{op} \triangleleft \bigwedge_{i \in I} \mu_i) \\ &= \bigwedge_{a \in X} R(x, a) * \bigwedge_{i \in I} (R^{op} \triangleleft \mu_i) \\ &\leq \bigwedge_{i \in I} \bigwedge_{a \in X} R(x, a) * (R^{op} \triangleleft \mu_i) \\ &= \bigwedge_{i \in I} (\alpha_R(\mu_i)(x)).\end{aligned}$$

□

Con el objeto de relacionar los lemas anteriores se componen los operadores apertura y erosión,

**Teorema 5.** Sea  $R$  un  $*$ -equivalencia se tiene que,

$$\alpha_R(\varepsilon_R(\mu)) \leq \varepsilon_R(\alpha_R(\mu))$$

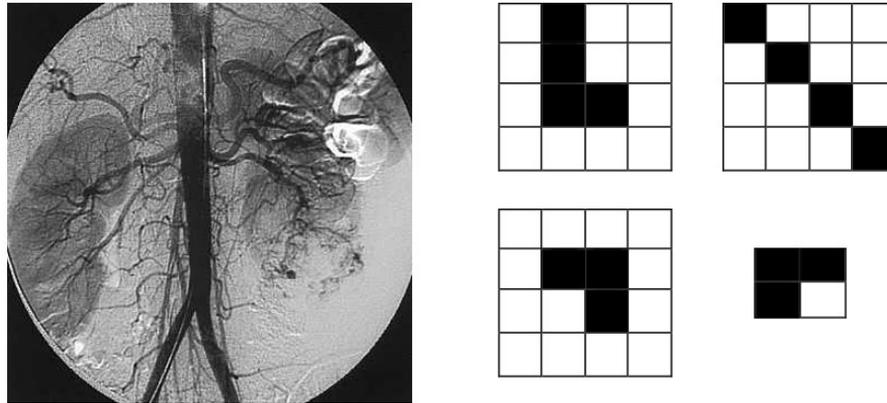
*Demostración.* Es evidente que  $\alpha_R(\varepsilon_R(\mu)) \leq \varepsilon_R(\mu)$  con  $\mu \in L^X$ , además se tiene,

$$\begin{aligned}\varepsilon_R(\mu) &= R^{op} \triangleleft \mu \\ &= R^{op} \triangleleft (R^{op} \triangleleft \mu) \\ &= R^{op} \triangleleft [R^{op} \triangleleft (R^{op} \triangleleft \mu)] \\ &\leq R^{op} \triangleleft [R^{op} \circ (R^{op} \triangleleft \mu)] \\ &= \alpha_R(\varepsilon_R(\mu))\end{aligned}$$

de lo anterior se tiene que,

$$\alpha_R(\varepsilon_R(\mu)) \leq \varepsilon_R(\alpha_R(\mu)).$$

□



**Figura 1.** Imagen original (arteriografía por sustracción digital con el método de Seldinger) y elementos estructurales.

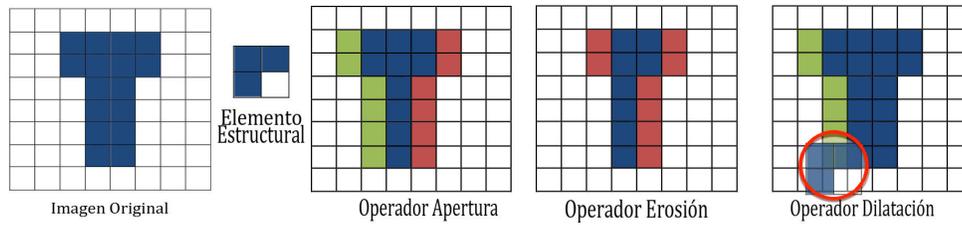
## 5. Aplicación

### 5.1. Materiales y Métodos

Un uso de los resultados presentados en la sección anterior se encuentra en el filtro de imágenes en escala de grises, lo cuál fue explorado por Forero en [10], en este caso se enfocará en imágenes médicas obtenidas por medio de contraste, la imagen que se implementará es una arteriografía por sustracción digital con el método de Seldinger tomada de la Figura 1.7 de [12] en formato JPEG, con dimensiones 1441 x 1441 píxeles, para ello se implementó el lenguaje Python y el paquete `pymorph` el cual permite aplicar los operadores erosión y apertura en el caso clásico, admitiendo un elemento estructural binario.

Con el propósito de entender la interpretación que realiza el computador de la imagen de estudio es necesario observar la Figura 2, en ella se tiene la implementación de los operadores, erosión, apertura y dilatación en el caso clásico, lo cual consiste en trasladar el elemento estructural a la imagen original, en la Figura 2, el elemento estructural (observe el círculo rojo en la Figura 2) se sitúa en cada píxel; se toma en el caso de la dilatación la intersección del elemento estructural con la imagen y en caso de que alguno de los píxeles este pintando, se colorea el píxel, en caso contrario se deja en blanco, una explicación mas detallada del caso clásico de la implementación de cada uno de los operadores antes mencionados en lenguaje Matlab - Python puede verse en [7].

En algunos casos es complicado determinar la diferencia entre las líneas sanguíneas y los órganos, para ello se debe tener un rango en el cual se pueda filtrar qué elementos de la imagen son vitales para el diagnóstico del médico, esto se logra por medio del paquete `OPENCV` en Python, específicamente el comando `inRange` el cual permite filtrar la imagen extrayendo los colores que se determinen; en este caso se optó por una escala de grises obtenidas con ayuda de la conversión de la imagen en formato HSV con el comando `COLOR_BGR2HSV` del mismo paquete, lo anterior simula la función de pertenencia del conjunto difuso que se está formando, cuyos elementos son



**Figura 2.** Aplicación de los operadores clásicos, verde ganancia, rojo perdida, azul valor original.

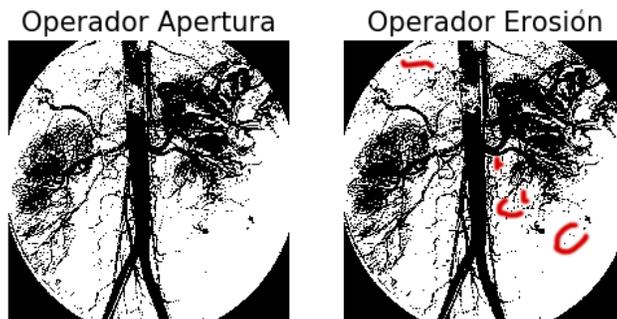
los píxeles de la imagen y su rango de pertenencia es obtenida tras la implementación del comando `inRange`.

Con los datos resultantes se procede a implementar el paquete `pymorph` y los códigos del operador apertura y erosión, con un elemento estructural generado por un arreglo de  $4 \times 4$  píxeles ajustado de forma adecuada para comprobar los resultados teóricos presentados en la sección anterior (se tomo el paquete `pymorph` sin modificar el código fuente debido a que la t-norma que implementa es la del mínimo).

## 5.2. Resultados

Se tiene presente que al filtrar imágenes de resonancia por medio de contraste del tórax con el paquete `pymorph` se debe tomar un elemento estructural adecuado. Si se quiere observar rastros sanguíneos es óptimo utilizar líneas dentro de un arreglo de píxeles.

Es oportuno preguntar qué ocurre si solo se tiene el caso clásico de blanco y negro o una tonalidad de gris y negro. La Figura 4 se construye al tomar un rango de pertenencia dada por una tonalidad clara (inferior izquierda), oscura (inferior derecha) y media (superior derecha) del color gris (en formato RGB son respectivamente, (211,211,211), (120,120,120) y (173,173,173)) y se aplica el operador apertura a cada una de las imágenes generadas al aplicar el rango de pertenencia a la ima-



**Figura 3.** El rojo indica ruido no eliminado y no el daño original en el sistema circulatorio del paciente.

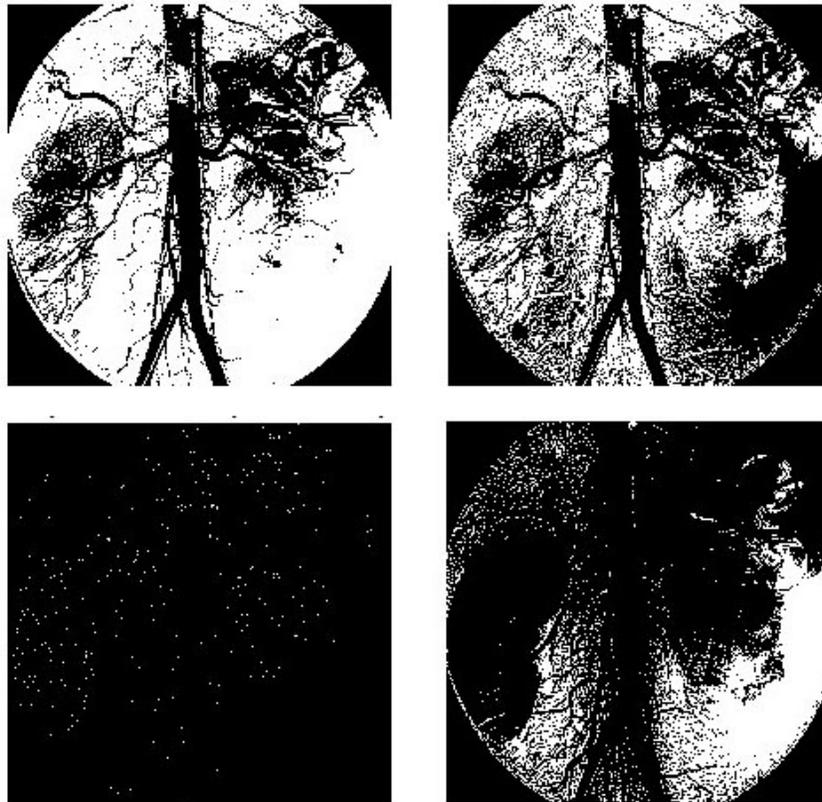


Figura 4. Operador apertura clásico aplicado con diferentes rangos de pertenencia.

gen original. Al unir estos intervalos de pertenencia en la escala (superior izquierda) se obtiene que la superposición de una con la otra da un mejor panorama de la imagen a analizar. Es así como se hace necesario una lógica de más de dos valores. Como se mostró en el sección anterior, si se aplica un número par de veces el operador apertura siempre que la relación estructural cumpla las condiciones del Teorema 4 es lo mismo que aplicarlo dos veces (debido a la idempotencia del operador), lo cual deviene en una reducción en las iteraciones que debe realizar el ordenador para arrojar la resonancia aplicando un filtro que genere como resultado tan solo las cavidades sanguíneas del paciente. En este caso se toma una relación estructural clásica, donde el elemento estructural es un cuadrado de  $2 \times 2$  píxeles. En la Figura 5, se exhibe la imagen original y el operador apertura aplicado una y diez veces.

Ahora, la Figura 6 se obtiene al filtrar con una relación (elemento estructural forma de L dentro de un cuadro de  $4 \times 4$  píxeles) que no satisface las condiciones del Teorema 4, para así resaltar la pérdida de información en el resultado que incide negativamente en el juicio del facultativo;

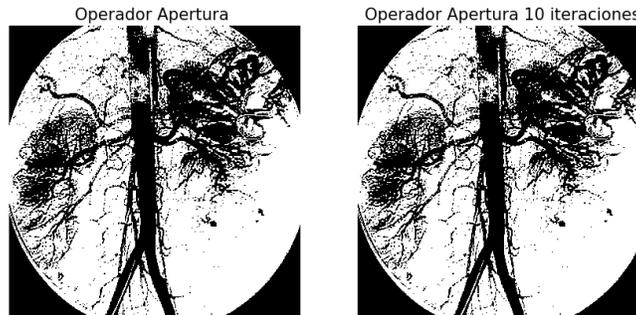


Figura 5. Operador apertura difuso con relación estructural reflexiva.

al evaluar que elementos son taponamientos y cuáles son naturales del sistema circulatorio del paciente; el ruido que aparece está resaltado en rojo.

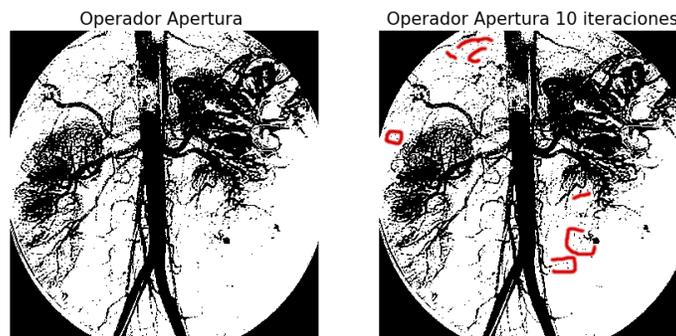
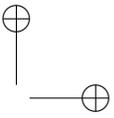
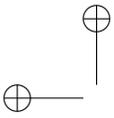


Figura 6. Operador apertura difuso con relación estructural no reflexiva ni \*-antitransitiva.

## 6. Discusión

En el estudio de imágenes diagnósticas es necesario la eliminación de ruidos para determinar anomalías del paciente, sea en la búsqueda de deformidades del feto como lo estudiaron A. P. Ballesteros, L. C. López, R. Herrera en [2] o taponamientos sanguíneos vistos en la sección anterior; esto último no es posible al aplicar los operadores morfológicos clásicos en Python por medio del paquete pymorph, debido a que admite tan solo un elemento estructural bivaluado. Es allí donde la contemplación de una relación estructural y la aplicación del rango de pertenencia de los valores a filtrar es necesario, en dicha búsqueda se encuentran los operadores morfológicos difusos y el intervalo de pertenencia se determina a partir del mínimo y máximo valor en la escala de grises de la imagen de estudio.

Además si se pregunta que operador aplicar a una imagen médica por medio de contraste en es-



cala de grises es pertinente recordar por lo demostrado en el Corolario 1 que el operador apertura poseerá mayor información que la erosión; y, si se tiene en cuenta una relación estructural adecuada, a su vez tendrá toda la información relevante con el menor ruido posible ya que podrá tener mayor información el operador apertura; pero si esta en su gran mayoría es ruido es infructuosa dicha información por lo tanto si se conjuga el Corolario 1 y el Teorema 4 se tendrá que la mejor opción es el operador apertura con una relación estructural bien sea reflexiva o \*-antitransitiva.

## 7. Conclusiones

La importancia de la matemática morfológica radica en el uso práctico que se le dé. Para mencionar solo un par de posibles aplicaciones en campos diferentes, ver por ejemplo su utilización en la detección de nódulos pulmonares en radiografías de tórax [1] o la identificación de lagunas y cuerpos de agua en imágenes geo-satelitales de alta resolución [16]. Un ejemplo mas concreto del uso de los operadores propuestos, se presentó en la sección 5 en imágenes en escala de grises muy usadas en diagnósticos de artereografía. Es por ello que si se tienen presentes los resultados en la elección de una relación estructural adecuada reducirá las iteraciones que deba realizar el computador para arrojar la imagen con la eliminación del ruido y perdida mínima de información relevante, además si se trabaja con el operador apertura es posible obtener menor ruido en la imagen que lo arrojado por la erosión.

Las implicaciones centrales en este artículo son:

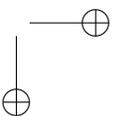
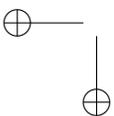
- Si  $R \in L^{X \times X}$  se tiene que para todo  $\mu \in L^X$ , se satisface que  $\varepsilon_R(\mu) \leq \alpha_R(\mu)$ .
- Si  $R$  es reflexiva y  $R^{op}$  es \*-antitransitiva entonces  $\varepsilon_R$  es un operador interior; además, si  $R$  es reflexiva o \*-antitransitiva entonces  $\alpha_R$  es un operador interior.
- Si  $R$  un \*-equivalencia entonces  $\alpha_R(\varepsilon_R(\mu)) \leq \varepsilon_R(\alpha_R(\mu))$ .

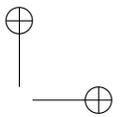
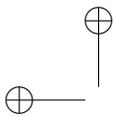
## Agradecimientos

Los aportes brindados por los evaluadores permitieron el crecimiento del presente artículo, por ese motivo extendemos nuestros más sinceros agradecimientos, además a los editores de la revista por su paciencia durante el proceso de publicación.

## Referencias

- [1] J. Aguillón, S. Duarte, R. Herrera, "Realce de candidatos a nódulo pulmonar en radiografías de tórax por medio de filtros de convergencia". *Revista de Ingeniería*, vol. 19, No. 2, pp. 85-104, 2014.
- [2] A. P. Ballesteros, L. C. López, R. Herrera, "Segmentación y conteo de las líneas de la nariz del feto en imágenes ecográficas de las 11-13+6 semanas de gestación". *Revista de Ingeniería*, vol. 20, No. 1, pp. 65-78, 2015.
- [3] R. Belohlávek, "Fuzzy relational systems, Ingeniería and Principles". *IFSR International Series on Systems Science and Engineering*, vol.20, Kluwer Academic/ Plenum Publishers, New York, 2002.
- [4] N. Bunce, R. Mohiaddin, "An Atlas of Contrast-Enhanced Angiography", *The Parthenon Publishing Group*, 2003.





- [5] R. Belohlávek, T. Funiokova, "Fuzzy Interior Operators", *International Journal of General Systems*, 2004.
- [6] N. Carmona, J. Elorza, J. Recasens, J. Bragard, "Permutable fuzzy consequence and interior operators and their connection with fuzzy relations". *Information Sciences*, Volumen 310, 20 July 2015.
- [7] R. Chityala, S. Pudipeddi, "Image Processing and Acquisition using Python". *Chapman Hall CRC Mathematical and Computational Imaging Sciences*, 2014.
- [8] E. Dougherty, R. Lotufo, "Hands on Morphological Image Processing". *SPIE Publications*, 2003.
- [9] J. Elorza, R. Fuentes-Gonzalez, J. Bragard, P. Burrillo, "One relation between fuzzy closing morphological operators, fuzzy consequence operators induced by fuzzy preorders and fuzzy closure and co-closure systems". *Fuzzy sets and systems*, pp. 73-89, May 2013.
- [10] W. Forero, "Relaciones difusas inducidas por el operador morfológico clausurativo difuso". Trabajo de Grado, Universidad Distrital Francisco José de Caldas, 2015.
- [11] U. Höhle, "Non-Classical Logics and their Applications to Fuzzy Subsets". *Springer Science+Business Media*, 1995.
- [12] A. R. Iturralde, "Urgencias urológicas". *Ciencias Médicas*, 2008.
- [13] J. Luna, C. Ochoa, "Interior Operators and Topological Categories". *Advances and Applications in Mathematical Sciences*, vol. 10, Issue 2, 2011, pp. 189-206, 2011.
- [14] D. Pei, "Fuzzy logic Algebras on Residuated Lattices". *Southeast Asian Bulletin of Mathematics*, 2004.
- [15] O. Salazar, J. Soriano, "Las leyes de tercero excluido y contradicción como valores límite en lógica difusa". *Revista de Ingeniería*, vol 16, No. 1, pp. 50-59, 2011.
- [16] G. Torrijos, "Extracción de cuerpo de agua utilizando técnicas de análisis de mezcla y morfología matemática". *Revista de Topografía Azimut*, vol 4, pp. 53-60, 2012.

---

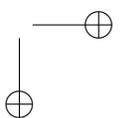
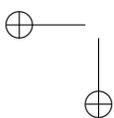
### **Carlos Orlando Ochoa Castillo**

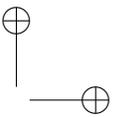
Licenciado en Educación con especialidad en Matemáticas y Física, Especialista en Matemática Aplicada y Magister en Ciencias Matemáticas de la Universidad Nacional de Colombia, Docente de la Universidad Distrital Francisco José de Caldas y Director del Semillero de Investigación LOTO.  
e-mail: karlos8ac@gmail.com

---

### **Wilson Javier Forero Baquero**

Estudiante de Maestría en Ciencias Matemáticas de la Universidad Nacional, Matemático de la Universidad Distrital Francisco José Caldas en la cuál fue líder del Semillero de Investigación LOTO.  
e-mail: wilsonforerob@gmail.com





Investigación

**Estudio Teórico Sobre la Influencia del Uso de Mezclas de Biodiesel de Aceite de Palma con Jet A-1 en Motores a Reacción**

*Theoretical Study of the Influence in a Jet Engine Using Mixtures of Jet A-1 and Biodiesel of Oil Palm*

**Luisa Fernanda Mónico Muñoz<sup>1</sup> , Juan José Sandoval Sotelo<sup>1</sup> , Andrés Felipe Rodríguez Chaparro\*<sup>1</sup>**

<sup>1</sup>Programa de Ingeniería Aeronáutica - Universidad de San Buenaventura. Sede Bogotá - Colombia.

Recibido: 05-09-2016. Modificado: 15-01-2017. Aceptado: 20-01-2016

Resumen

**Contexto:** La industria aeronáutica, al igual que muchas otras, se ha visto obligada a implementar en la actualidad métodos para mitigar el daño producido al medio ambiente por la emisión de gases contaminantes y, de esta manera, hacerle frente a la problemática del calentamiento global. En este panorama, la investigación sobre la utilización de combustibles alternativos es de suma relevancia, en particular el estudio del rendimiento de motores al usar mezclas de biodiésel colombiano a base de aceite de palma con Jet A-1.

**Método:** Como punto de partida, se realiza una revisión al estado del arte para la selección del motor a estudiar y los combustibles alternativos a utilizar; paralelamente, se seleccionó el biodiésel colombiano a base de aceite de palma, ya que ha demostrado tener un buen comportamiento como combustible al ser mezclado en porcentajes de 10 %, 20 % y 50 %. A cada una de las mezclas realizadas se les hicieron pruebas en laboratorio para conocer su valor de viscosidad, densidad y poder calorífico; para la realización de los análisis, se desarrolló un programa sobre una hoja de cálculo que contempla propiedades físicas de las mezclas y parámetros del motor, así como emisiones de óxidos nitrosos (NOx), atomización y combustión producida por el uso de las mezclas. Los resultados se contrastaron contra el software de simulación Gas turbine.

**Resultados:** Se escogió el CFM56-5B como motor apropiado para realizar el estudio debido a que, en la actualidad, es el más empleado en la industria aeronáutica colombiana; por otra parte, ninguna de las mezclas de combustible alternativo genera el mismo rendimiento del motor al emplear Jet A-1. Las mezclas E10 y E20 tienen un comportamiento similar, con la novedad de generar menor cantidad de emisiones de NOx y mejorar la atomización del combustible.

**Conclusiones:** Si se compara el rendimiento del motor al usar combustible convencional y al emplear las mezclas de biodiésel, se observa una disminución del empuje conforme aumenta el porcentaje de biodiésel en la mezcla; debido a que las mezclas de combustible alternativo tienen un menor poder calorífico, el consumo de combustible en estos casos es siempre mayor. Por otro lado, al comparar los niveles de NOx producidos, se aprecia una disminución de este contaminante al usar las mezclas de biodiésel, vale la pena resaltar que es de gran interés para trabajos futuros evaluar otras emisiones contaminantes y, a su vez, el comportamiento de otros tipos de biodiésel en motores a reacción.

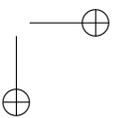
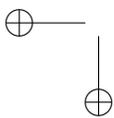
**Palabras clave:** Combustible alternativo , emisiones contaminantes, motor a reacción, prestaciones.

**Idioma:** Español



Citación: L. Mónico, J. Sandoval, A. Rodríguez. "Estudio Teórico Sobre la Influencia del Uso de Mezclas de Biodiesel de Aceite de Palma con Jet A-1 en Motores a Reacción". INGENIERÍA, vol. 22, no. 1, pp. 140-151, 2017.

©Los autores; titular de derechos de reproducción Universidad Distrital Francisco José de Caldas. En línea DOI: <http://dx.doi.org/10.14483/udistrital.jour.reveng.2017.1.a09>



### Abstract

**Context:** Like many others, today the aeronautical industry has been forced to implement methods to mitigate the damage produced to the environment due to the emission of polluting gases and in doing so, confront the problem of global warming. In this context, research on the use of alternative fuels is of paramount importance, in particular the study of engine performance when using blends of Colombian Biodiesel based on palm oil with Jet A1.

**Method:** As a starting point, we made a review of the state of the art so as to select which engine to study and then the alternative fuels to be used. Simultaneously, Colombian biodiesel based on palm oil was selected, as it has been shown to be a fuel with good performance when it is blended in percentages of 10 %, 20 % and 50 %. Each of the blends were tested in the laboratory to obtain their viscosity, density and calorific values. A spreadsheet program was developed to conduct the analysis, which contemplated the physical properties of mixtures and engine parameters, as well as emissions of nitrous oxides (NO<sub>x</sub>), atomization and combustion produced by the use of mixtures. The results were compared against those of the Gasturbine simulation software.

**Results:** CFM 56-5B was chosen as the appropriate engine for the study, because it is currently the most used in the Colombian aeronautical industry. On the other hand, none of the alternative fuel blends generates the same engine performance when using Jet A1. The blends E10 and E20 have a similar behavior, with the novelty of generating less amount of NO<sub>x</sub> emissions and improving the atomization of the fuel.

**Conclusions:** When the engine's performance with conventional fuel is comparing with Biodiesel mixtures, a decrease of the thrust is produced as the percentage of Biodiesel in the mixture increases. Because the mixtures of alternative fuel have a lower calorific value fuel consumption in these cases is always greater. However, when comparing the NO<sub>x</sub> levels produced, a decrease of this pollutant is observed when using Biodiesel blends. It is worth noting that it is of great interest for future work to evaluate other pollutant emissions and, in turn, the behavior of other types of Biodiesel in jet engines.

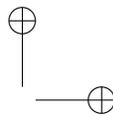
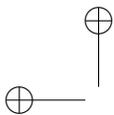
**Keywords:** Alternative fuel, engine, performance, pollutant emissions.

## 1. Introducción

En los últimos años se ha venido experimentando mayor evidencia de la problemática ambiental conocida como calentamiento global, la cual surge, principalmente, por las emisiones no controladas de gases efecto invernadero que son producidas en su mayoría por las industrias químicas y de transporte. La industria aeronáutica es la responsable del 2 % de las emisiones de CO<sub>2</sub>. [1], es por esto que la Asociación Internacional de Transporte Aéreo (IATA) se comprometió a “mejorar la eficiencia del consumo de combustible un 1.5 % anual hasta 2020, estabilizar las emisiones de CO<sub>2</sub> con un crecimiento neutro a partir de 2020 y reducir las emisiones hasta el 50 % en 2050” [1].

De esta manera, mediante el trabajo conjunto entre universidades de todo el mundo y la industria aeronáutica, se han realizado diferentes investigaciones en temas de disminución de consumo específico de combustible (SFC) en motores a reacción, mejoras en los procesos de combustión, reducción de óxidos nitrosos (NO<sub>x</sub>) y dióxido de carbono (CO<sub>2</sub>) mediante la implementación de combustibles alternativos [2]–[7].

Los primeros avances fueron realizados por Continental Airlines, quien en el año 2009 logró rea-



lizar el primer vuelo comercial en un avión Boeing 737 utilizando una mezcla de 50 % biocombustible a base de algas y 50 % de Jet A-1 en uno de sus motores [8]. Posterior a este vuelo, se siguieron realizando pruebas para determinar qué combustible alternativo podría llegar a suplir el combustible convencional, sin necesitar un cambio en la estructura y diseño del motor, ya que los motores no pueden ser fácilmente modificados en todas las flotas; pensando en esto, se determinó que los combustibles que se deseen emplear en la aviación deben cumplir con los requerimientos del motor, brindando prestaciones parecidas al motor al emplear un combustible convencional [1].

A raíz de lo anterior, entre los años 2009 y 2011 el estudio de mezclas con diferentes combustibles alternativos, al igual que las pruebas experimentales, se hicieron más frecuentemente en vuelos comerciales [8], [9]; además, en el 2011, la armada de los Estados Unidos realizó una demostración en motores turbo eje, utilizando una mezcla 50 % de combustible alternativo solazyne y 50 % de combustible convencional, en un helicóptero MH-60S Seahawk, marcando el inicio de las pruebas con biocombustibles en los motores a reacción de vehículos aéreos militares [9].

Por otro lado, enfocándose en la industria colombiana, hay un potencial para ser pioneros en la implementación de combustibles alternativos en motores de aviación, ya que el país produce combustibles como el biodiésel a partir del aceite de palma y el bioetanol derivado de la caña de azúcar [2], [3]; en estos, basándose en las pruebas realizadas en otros lugares del mundo y conociendo que gran parte de los ensayos se hicieron utilizando diferentes tipos de biodiésel, se busca aprovechar el auge y la disponibilidad de los combustibles alternativos en Colombia para estudiar cuál es su influencia en un motor a reacción, comúnmente usado en la industria, bajo la influencia de mezclas realizadas entre los combustibles alternativos disponibles (aptos para este tipo de motores) y el Jet A-1.

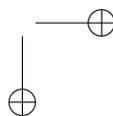
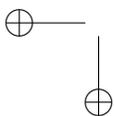
## 2. Metodología

Inicialmente, se realiza un estado del arte para la selección de los combustibles alternativos y el tipo de motor a utilizar, esto basado en resultados publicados en artículos de investigación; de esta manera, se determinan cuáles combustibles alternativos, al mezclarse con combustible tradicional, muestran buen desempeño en motores a reacción en términos de actuación, atomización, eficiencia en la combustión y reducción de las emisiones contaminantes.

Sin embargo, muchos de estos combustibles no se encontraban disponibles en la industria colombiana, un ejemplo es el combustible derivado de la jatropha o el biodiésel de algas [9]; el biodiésel de aceite de palma, por otro lado, demostró tener un buen comportamiento (relacionado con las prestaciones) como combustible en motores al ser mezclado en porcentajes de 10 %, 20 % y 50 % [9].

Posteriormente, mediante la prueba ASTM D240 realizada en la Universidad Nacional de Colombia, fue posible determinar el poder calorífico de las mezclas; así mismo, en los laboratorios de la Universidad de San Buenaventura, sede Bogotá, empleando la norma ASTM D1217 se halló la densidad y con la Ecuación 1 se determinó la viscosidad.

$$\mu_m = \mu_1^{x_1} * \mu_2^{x_2} \quad (1)$$



Los resultados de las pruebas (ver Figura 1) muestran que la viscosidad de la mezcla aumenta al tener mayor porcentaje de biodiésel, este comportamiento es lógico debido a que el biodiésel es más viscoso que el Jet A-1.

Así mismo, a partir de la Figura 2 se puede observar que la mezcla con mayor poder calorífico es la E10, seguida de la E20 y por última la mezcla E50; el comportamiento que se observa del poder calorífico de las mezclas es el resultado de agregar mayor porcentaje de biodiésel, el cual tiene menor poder calorífico. Esta disminución respecto al poder calorífico del Jet A-1

De igual forma, es importante observar que, en la Figura 3, hay un cambio lineal en las densidades, es decir, a medida que se va agregando mayor cantidad de biodiésel a la mezcla, esta adquiere un mayor valor de densidad y viceversa; este comportamiento es de esperarse, ya que el biodiésel es más denso y si se aumenta la cantidad que se usa en la mezcla se generará un aumento de esta propiedad.

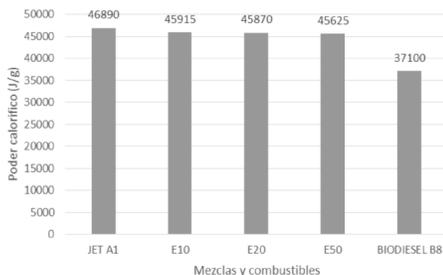


Figura 2. Poder calorífico de las mezclas.

De forma paralela a la realización de las pruebas, se seleccionó el motor CFM56-5B debido a que en la industria colombiana existe una predominancia en la familia Airbus A320, los cuales usan este tipo de motores; además, los datos generales del diseño se encontraban disponibles para su uso en el estudio. Los parámetros generales del motor se pueden observar a continuación en la Tabla I [9].

Tras la selección del motor, el proceso se complementó con dos tareas: en primer lugar, se ingresaron los datos de las propiedades físicas de

es tan solo del 2.1 % para el E10, 2.2 % para el E20 y 2.7 % para el E50.

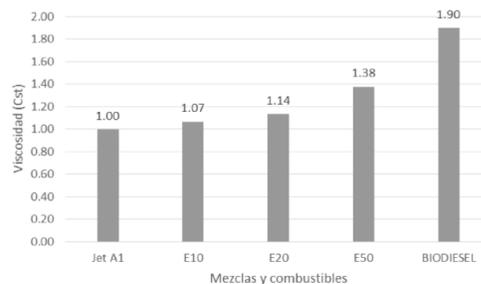


Figura 1. Viscosidad de las mezclas.

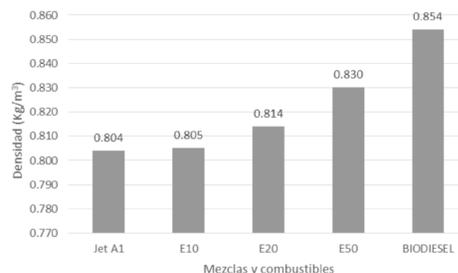


Figura 3. Densidad de las mezclas.

Tabla I. Parámetros generales del motor

Especificaciones	Valor
Empuje	98 kN
Tipo de motor	Turbofan, flujo axial
Tipo de cámara de combustión	Anular
Tipo compresor	Doble eje, 13 etapas (4 baja y 9 alta)
Tipo de turbina	5 etapas (1 alta y 4 baja)
Peso del motor	1946 kg en seco
Largo del motor instalado	489.2 cm
Diámetro motor	224.5 cm
Sistema de ignición	2 por cada motor
Aceite lubricante	Mobil Jet Oil II
Combustible	JET-A1

las propiedades físicas de las mezclas y los parámetros del motor como la relación de compresión, temperatura de entrada, entre otros, a la programación realizada por los autores en una hoja de cálculo de Excel, basándose en la teoría del libro *Gas turbine theory* [10], para así determinar la afectación del motor a reacción seleccionado. Posteriormente se comprobaron dichos resultados mediante la utilización del *software* de simulación Gas turbine; para finalizar, se añadió a la programación hecha los cálculos de las emisiones de  $\text{NO}_x$ , el cambio en la atomización y en la combustión que produce el uso de las mezclas seleccionadas en este artículo.

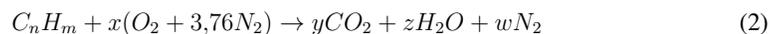
### 3. Resultados

#### 3.1. Rendimiento del motor

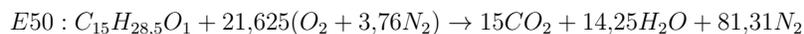
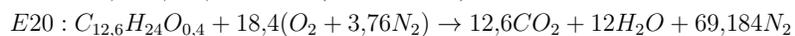
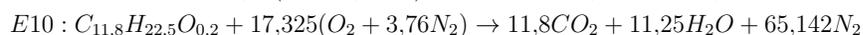
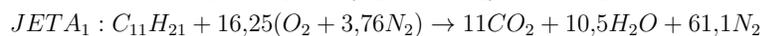
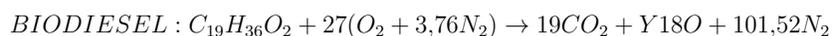
Para determinar la afectación del rendimiento del motor, se hizo un análisis suponiendo una entrada constante de flujo másico de combustible de 3 kg/s (flujo aproximado real del motor usando Jet A-1 brindado por la empresa International CFM).

Como punto de partida, fue necesario hallar la relación combustible/aire (C/A) de cada una de las mezclas, se conoce que la fórmula aproximada del Jet A-1 es  $\text{C}_{11}\text{H}_{21}$  y del biodiésel es  $\text{C}_{19}\text{H}_{36}\text{O}_2$  [11]; posteriormente, se hizo una interpolación para determinar la fórmula preliminar de cada mezcla, este procedimiento consiste en tomar el porcentaje de la mezcla del biodiésel en número decimal y multiplicarlo por el número correspondiente al subíndice de la molécula de la fórmula química del biodiésel. De igual forma se realizó con el Jet A-1.

Más adelante, se hizo el balanceo de ecuaciones mediante el método de coeficientes indeterminados, mostrado como ejemplo a continuación:



Se muestran a continuación las fórmulas químicas balanceadas para cada una de las mezclas:



Luego del balanceo de las ecuaciones químicas, se usa el peso atómico de cada componente del aire ( $\text{O}_2 + 3,76\text{N}_2$ ) multiplicado por el valor hallado de  $X$ , esto se divide por la multiplicación de los pesos atómicos del combustible con sus respectivos subíndices. De los resultados obtenidos, se puede observar en la Figura 4 que a medida que la mezcla contiene mayor cantidad de biodiésel, la relación combustible/aire aumenta; este comportamiento está directamente relacionado con los subíndices de la fórmula química del biodiésel, los cuales son mayores que los del Jet A-1, seguido por el E10, E20 y E50 [12].

Posteriormente, introduciendo los datos en la programación realizada y empleando información específica del Manual de mantenimiento del motor CFM56-5B, se calcula el empuje (Ecuación 3) y el consumo específico del combustible respecto al empuje (TSFC) (Ecuación 4).

$$T = \dot{m}_e V_e - \dot{m}_0 V_0 + BPR \dot{m}_c V_F \dot{m} \quad (3)$$

$$TSFC = \frac{SFC}{T} \quad (4)$$

A partir de la Figura 5, es posible observar que el empuje del motor disminuye a medida que la mezcla contiene mayor cantidad de biodiésel; esta reducción en el empuje se debe a que la relación combustible/aire cambia, por lo tanto, respecto al empuje del motor con Jet A-1, se disminuye para el E10, E20, E50 y biodiésel en un 2.42 %, 4.6 %, 9.26 % y 14.14 % respectivamente.

Por otro lado, de acuerdo al consumo específico de combustible respecto al empuje, se obtiene que debido al aumento en la relación combustible/aire para cada mezcla hay un incremento en el consumo; este comportamiento se debe a que las mezclas, entre mayor cantidad de biodiésel contengan, requerirán de una mayor cantidad de combustible para igualar el empuje generado al emplear el combustible convencional.

Así mismo, se puede observar en la Figura 6 que la mezcla con mayor aumento en el consumo específico de combustible respecto al empuje (TSFC) es la E50, la cual aumentó, respecto al valor calculado del Jet A-1, en 9.74 % su consumo; de igual manera, el aumento en este parámetro es del 4.69 % para el E20, y 2.48 % para el E10.

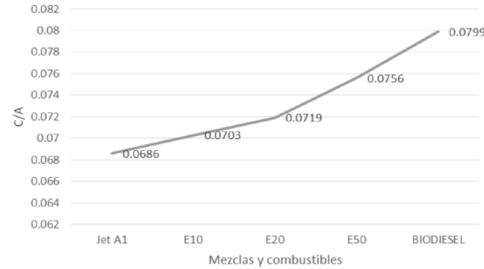


Figura 4. Relación combustible/aire.

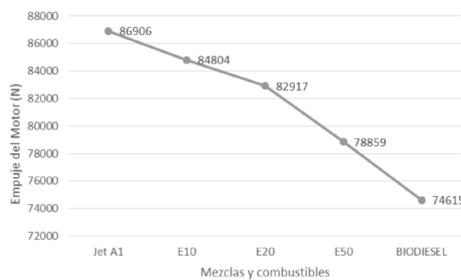


Figura 5. Empuje del motor usando mezclas de combustible.

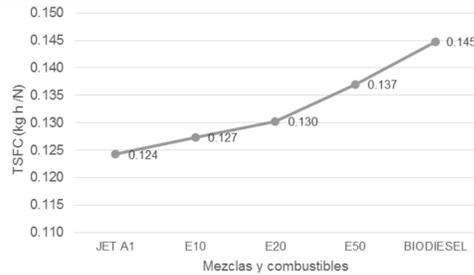


Figura 6. TSFC.

### 3.1.1. Comparación con los resultados de Gasturb

Después de ingresar todos los datos necesarios para la simulación en el programa Gasturb, lo ideal es que el empuje sea el mismo o se obtenga un margen de error bajo; en la Tabla II, se muestran los resultados de la comparación del empuje obtenido en la programación hecha por los autores y la

la calculada por el programa Gasturb, en dicha tabla se observa que el porcentaje de error en todos los casos es menor al 4 %, lo cual comprueba que la metodología usada es correcta.

**Tabla II.** Porcentaje de error.

Combustible	Empuje obtenido en programación (N)	Empuje obtenido en simulación Gasturb (N)	% error
Jet A-1	86.906	83.85	3.5
E10	84.804	81.84	3.4
E20	82.917	80.19	3.2
E50	78.859	76.15	3.4
Biodiésel	74.615	72.87	2.3

### 3.2. Influencia en la atomización

Para determinar la influencia de las mezclas en la atomización se hace uso de la Figura 7, la cual se caracteriza por clasificar los chorros de salida basándose en el número de Ohnesorge (Ecuación 5), el número de Reynolds (Ecuación 6), el número de Wobbe (Ecuación 7) y la densidad de energía (Ecuación 8), para determinar hasta qué punto es posible el intercambio de los combustibles.

$$Oh = \mu_L / (\rho_L \sigma D)^{0.5} \quad (5)$$

$$Re = \frac{\rho v D}{\mu} \quad (6)$$

$$We = \frac{Pc}{\sqrt{\frac{\rho_{comb}}{\rho_{aire}}}} \quad (7)$$

$$DE = Pc \rho \quad (8)$$

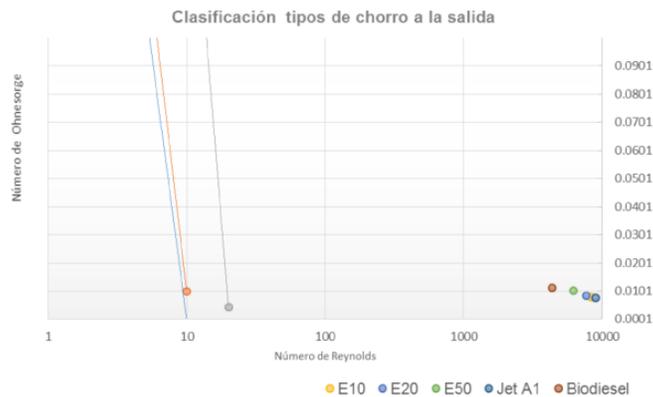
De esta manera, previamente se ingresan los datos de cada mezcla a la programación para obtener los resultados mostrados en la Tabla III.

**Tabla III.** Numero de Reynolds y número de Ohnesorge

Combustible	Número de Reynolds	Número de Ohnesorge
E10	8297398	0.0066
E20	7684056	0.0070
E50	6183537	0.0083
JET A-1	8907925	0.0062

Luego, al ubicar los valores de la Tabla III en la Figura 7, se puede observar bajo qué régimen de inyección está ubicada cada una de las mezclas para las configuraciones del motor.

De acuerdo a la Figura 7 y la Tabla III, se observa que, al agregar mayor cantidad de biodiésel a la mezcla, el número de Ohnesorge aumenta, ubicando las mezclas con mayor cantidad de biodiésel



**Figura 7.** Clasificación tipos de chorro a la salida utilizando E10, E20, E50 Y JET A1.

en las partes más altas de la gráfica; no obstante, todas las mezclas permanecen dentro de la zona de atomización.

Conforme a lo anterior, se encuentra que todas las mezclas se atomizan en forma de gotas diminutas, generando el comportamiento necesario para que se dé una combustión eficiente. Dicho comportamiento se debe a que el biodiésel tiene mayores valores de densidad y viscosidad.

Por otro lado, al calcular el índice de Wobbe y la densidad energética, se obtiene la Tabla IV.

**Tabla IV.** Índice de Wobbe y Densidad energética

<b>Índice de Wobbe (J)</b>				
<b>JET A-1</b>	<b>E10</b>	<b>E20</b>	<b>E50</b>	<b>BIODIESEL</b>
59459.04	58186.51	57807.24	56941.58	45646.83
-	2 %	3 %	4 %	23 %
<b>Densidad energética (J/m )</b>				
<b>Jet A-1</b>	<b>E10</b>	<b>E20</b>	<b>E50</b>	<b>BIODIESEL</b>
37699560	36961575	37338180	37868750	31683400
-	2 %	1 %	-0.4 %	16 %

En estos resultados se puede observar que la intercambiabilidad de combustible es posible para los combustibles E10, E20 y E50, ya que el porcentaje de disminución con respecto al Jet A-1 es menor al 5 % en ambos casos, e incluso en el caso del E50 se tiene un aumento en la densidad energética, lo cual significa que, si los combustibles son almacenados en un tanque de volumen determinado, será la que mayor energía podrá entregar, seguida del Jet A-1, E20 y E10; en el caso del biodiésel, como resulta evidente, la intercambiabilidad no es recomendada dada la alta disminución porcentual del índice de Wobbe y la disminución energética en un mismo contenedor.

### 3.3. Influencia en la combustión

Para hallar la eficiencia en la combustión de las mezclas se usa la Ecuación 9 [11].

$$\eta_{cc} = \frac{\xi * G_{Hora} * P_c}{V_{cc} * P_A} \quad (9)$$

Por lo tanto, al usar los valores presentados de poder calorífico para cada mezcla, y la presión a la entrada de la cámara de combustión de

3350000 Pa obtenida en la programación, se obtienen los valores presentados en la Tabla V.

**Tabla V.** Rendimiento de la combustión

<b>Combustible</b>	<b>Pc (J/g)</b>	$\eta_{cc}$	<b>Error relativo</b>
JET A-1	46890	0.99	
E10	45915	0.96	2.08 %
E20	45870	0.96	2.18 %
E50	45625	0.96	2.70 %
BIODIESEL B8	37100	0.78	20.88 %

Analizando los resultados obtenidos en esta tabla, se puede relacionar que a medida que el poder calorífico disminuye, el rendimiento en la combustión también; esto ocurre porque todos los valores usados en la Ecuación 2, a excepción del poder calorífico, permanecen constantes (la cámara de combustión no cambia y el flujo másico de combustible para este análisis tampoco).

Por lo tanto, como se indicó en las pruebas de poder calorífico, el biodiésel es el combustible

con menor poder calorífico, lo cual produce una reducción en la eficiencia de la combustión del 20.88 % respecto al Jet A-1, seguido del E50 con reducción en el 2.70 %, E20 con 2.18 % y E10 con reducción de tan solo 2.08 %.

### 3.4. Resultados emisiones de NO<sub>x</sub>

Inicialmente, es necesario calcular o usar datos experimentales de la temperatura de la llama, en este caso llama adiabática; sin embargo, si se hace de manera teórica, puede llevar a tener mayores errores debido a múltiples asunciones, como las entalpías de reacción no determinadas de manera experimental del biodiésel y de las mezclas.

Por lo tanto, de acuerdo a una investigación hecha por la Universidad de Cambridge, se determinó de manera experimental la temperatura de llama adiabática del biodiésel derivado de aceite de palma y del Jet A-1 [14]; de esta manera, extrayendo la información del documento nombrado, se sabe que la temperatura de llama adiabática del Jet A-1 es de 2587 K y del biodiésel es de 2564 K.

Posteriormente, haciendo una interpolación respecto al porcentaje de la mezcla, el cual consiste en tomar el porcentaje del combustible y multiplicarlo por el respectivo valor de la temperatura adiabática, se determina la temperatura de la llama adiabática para cada mezcla, como se muestra en la Figura 8.

La relación entre la temperatura de la llama y el índice de emisión de NO<sub>x</sub> para motores turbo-reactores es analizada empleando la Ecuación 10 [15].

$$EINO_x = 2 * 10^{-83} * T_{flama}^{24,696} \tag{10}$$

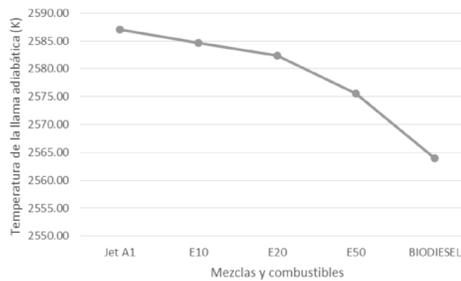


Figura 8. Temperatura de la llama adiabática.

En la Tabla VI, se muestran los resultados de la temperatura de llama adiabática e índice de emisiones de NO<sub>x</sub>, para cada uno de los combustibles y mezclas analizados.

**Tabla VI.** Comparación índice de NO<sub>x</sub>

	Jet A-1	E10	E20	E50	BIODIESEL
Temperatura adiabática (K)	2587	2584.7	2582.4	2575.5	2564
EI NO <sub>x</sub> (g/kg)	38.32	37.49	36.67	34.33	30.74
Porcentaje de disminución	-	2 %	4 %	10 %	20 %

A partir de los datos que se relacionan en la Tabla VI, se deduce que a medida que la mezcla contiene mayor cantidad de biodiésel, la temperatura de llama adiabática es menor, generando que las emisiones por NO X disminuyan en un 2 %, 4 %, 10 % y 20 % para las mezclas E10, E20, E50 y el combustible biodiésel B8, respectivamente. Este comportamiento es definido por la naturaleza de los componentes de los cuales está formado cada uno de los combustibles, relacionado por las entalpías de formación.

## 4. Conclusiones

Conforme se determinaron los valores de las propiedades físicas de la viscosidad, densidad y poder calorífico para cada una de las mezclas, se observa que a medida que se le agrega mayor cantidad de biodiésel a la mezcla el poder calorífico de esta disminuye, pero la viscosidad y la densidad aumentan; esto ocurre porque el biodiésel contiene menor poder calorífico pero mayor viscosidad y densidad. Por lo tanto, para cargar el mismo volumen de biodiésel que el Jet A-1 en un Airbus A320, el cual tiene tanques de combustible con capacidad de hasta 19287.607 m<sup>3</sup>, aumentaría el peso en 963.8 kg (aproximadamente 6 pasajeros con maletas), de igual manera para el E50, E20 y E10, el peso tendría un aumento de 500.192 m<sup>3</sup> y 19.2 kg respectivamente.

A partir del balanceo de las ecuaciones, se encuentran que la relación combustible/aire disminuye en 16 %, 10 %, 5 % y 3 % para el biodiésel, E50, E20 y E10 respectivamente, generando que se requiera mayor cantidad de combustible para mantener el balanceo estequiométrico; sin embargo, no afecta en porcentajes importantes las mezclas E10 y E20.

El mayor empuje del motor se obtiene al usar Jet A-1, debido a que tiene un mayor poder calorífico; por otro lado, para la mezcla E10, E20 y E50 el empuje se disminuye en 2.42 %, 4.6 % y 9.26 % respectivamente, es decir que, para la parte del rendimiento, continúa siendo mejor usar Jet A-1.

Por otro lado, de acuerdo al consumo específico de combustible respecto al empuje, se obtiene que debido al aumento en la relación combustible/aire para cada mezcla hay un incremento en el consumo; este comportamiento se debe a que entre mayor cantidad de biodiésel contengan las mezclas menor empuje generan, además requieren de mayor cantidad de combustible. Así, el aumento en el consumo de combustible aumenta en un 2.48 %, 4.69 %, 9.74 % y 14.95 %, para el E10, E20, E50 y el biodiésel respectivamente, comparados con el Jet A-1.

De acuerdo a los resultados de la sección de atomización, se puede concluir que todas las mezclas y los combustibles se mantienen en la zona denominada “atomización”, es decir, se vuelven gotas en zonas de la tobera de salida, siendo un comportamiento adecuado; sin embargo, el biodiésel sin mezclar es el que tiene mejor características de atomización, seguido del E50, E20 y E10, lo cual implica que a mayor cantidad de biodiésel que tenga la mezcla, tendrá mejor eficiencia en la atomización.

En el caso de la intercambiabilidad de combustibles se encuentra que las tres mezclas obtienen muy buenos resultados, ya que los porcentajes de disminución son por debajo del 4 %, e incluso en la mezcla E50 se tiene una mayor densidad energética, lo cual generaría que se conservase el buen rendimiento al usar estas mezclas.

Dado que la eficiencia en la combustión con una misma cámara de combustión varía de acuerdo al poder calorífico del combustible y este disminuye a medida que la mezcla contiene mayor cantidad de biodiésel, hay una reducción de la eficiencia de las mezclas respecto al Jet A-1 del 3 %.

Se determina que hay una disminución del 20 % del índice de NO<sub>x</sub> cuando se usa biodiésel en comparación al Jet A-1; de igual forma para las mezclas E50, E20 y E10, se disminuye este índice en 10 %, 4 % y 2 % respectivamente.

## Referencias

- [1] IATA. *Trabajando juntos en materia de combustible –los biocombustibles son una prioridad para la industria aeronáutica*. 2011. [En línea] Disponible en: <https://www.iata.org/pressroom/pr/Documents/Spanish-PR-2011-11-15-01.pdf> ↑ 141, 142
- [2] Barón, M., Huertas, I. y Orjuela, J., “Gestión de la cadena de abastecimiento del biodiésel: una revisión de la literatura”. *Revista de Ingeniería*, Vol 18, No 1., pp. 84-117, 2013. ↑ 141, 142
- [3] ARanda, J., Barón, M., Forero, I., Orjuela, J., “Modelo de programación matemática para la cadena productiva del biodiesel en Colombia”. *Revista de Ingeniería* Vol 19, No 1., pp. 19-49, 2014. ↑ 141, 142
- [4] Aranda, J., Orjuela, J., “Optimización multiobjetivo en la gestión de cadenas de suministro de biocombustibles. Una revisión de la literatura”. *Revista de Ingeniería* Vol. 20, No 1., pp. 37-63, 2015. ↑ 141
- [5] lanjekar, R.D. Deshmukh, D. “A review of the effect of the composition of biodiesel on NO<sub>x</sub> emission, oxidative stability and cold flow properties”. *Renewable and Sustainable Energy Reviews*, Vol. 54. pp. 1401-1411, 2016. ↑ 141
- [6] Badami, M. Nuccio, P. Pastrone, D. Signorettoa, A., “Performance of a small-scale turbojet engine fed with traditional and alternative fuels”. *Energy Conversion and Management*, Vol. 82. pp. 219-228, 2014. ↑ 141
- [7] Natelson, R. H. Wang, W. Roberts, W. L. Zering, K. D. “Technoeconomic analysis of jet fuel production from hydrolysis, decarboxylation, and reforming of camelina oil”. *Biomass and Bioenergy*, Vol. 75. pp. 23-34, 2015. ↑ 141
- [8] Murrilo, J., *Producción de biodiesel a partir de aceite de palma*. Universidad Nacional de Manizales, pp. 13-16, 2003. ↑ 142
- [9] Allaboutalgae, *Renewable Aviation Fuel*, 2015. [En línea] Disponible en: <http://allaboutalgae.com/fuel-tests/> ↑ 142, 143
- [10] Saravanamuttoo, H.I.H. Rogers, G.F.C. Cohen, H. Straznycky, P. *Gas turbine Theory* [ed.] Cuarta, Pearson, Inglaterra, pp. 336-369, 1996. ↑ 144
- [11] Oñate, A. E. *Turborreactores teoría, sistemas y propulsión de aviones*, Ed. Sumaas, S.A / A., 1981. ↑ 144, 147
- [12] Martínez, I. *Combustion Kinetics*, 1995. [En línea] Disponible en: <http://websserver.dmt.upm.cs/isidoro/bk3/cis/Combustion%20Kinetics.pdf> ↑ 144
- [13] Tung, C. Hochgreb, S. *Measurements of laminar flame speeds of liquid fuels: Jet- A1, diesel, palm methyl esters and blends using particle imaging velocimetry* (PIV, ed. 33), Combustion Institute, pp. 979–986, 2011. ↑
- [14] Glaude, R. Fournet, R. Bounaceur, And M. Molière, “Adiabatic flame temperature from biofuels and fossil fuels and derived effect on NO<sub>x</sub> emissions”. *Fuel Proc. Technol*, vol. 91, pp. 229-235, 2010. ↑ 148
- [15] Pearce, G.F., Kretschmer, D., Odgers, J., Wang, G. “The prediction of thermal NO<sub>x</sub> in gas turbine exhausts”, in: *Eleventh International Symposium on Air Breathing Engines*, Tokyo, 1993, pp. 1–9. ↑ 148

---

### Luisa Fernanda Mónico Muñoz

Nació en Bogotá, Colombia. Es Ingeniera Aeronáutica de la Universidad San Buenaventura, de Bogotá, Colombia. En el 2010, obtuvo su título de Maestría en Motores de Combustión Interna Alternativos en la Universidad Politécnica de Valencia de Valencia, España. En el 2013 terminó su doctorado en Sistemas Propulsivos en Medios de Transporte en la Universidad Politécnica de Valencia de Valencia, España. Actualmente se desempeña como docente del Programa de Ingeniería Aeronáutica en la Universidad de San Buenaventura, sede Bogotá, Colombia, y pertenece al grupo de investigación Aerotech donde realiza estudios en la línea de Energía y Propulsión. e-mail: [lmonico@usbog.edu.co](mailto:lmonico@usbog.edu.co)

---

### Andrés Felipe Rodríguez Chaparro

Nació en Bogotá, Colombia. Ingeniero Aeronáutico de la Universidad San Buenaventura, de Bogotá, Colombia. e-mail: [afrodriguez@academia.usbbog.edu.co](mailto:afrodriguez@academia.usbbog.edu.co)

---

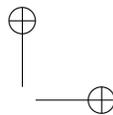
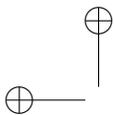
### Juan José Sandoval Sotelo

Nació en Cartagena, Colombia. Ingeniero Aeronáutico de la Universidad San Buenaventura, de Bogotá, Colombia. e-mail: [jjssandoval@academia.usbbog.edu.co](mailto:jjssandoval@academia.usbbog.edu.co)

## Nomenclatura y símbolos

Nomenclatura	Significado	Unidad
BPR	By Pass Ratio	
CO <sub>2</sub>	Dióxido de carbono	
D	Diámetro de la gota	m
DE	Densidad energética	J/m <sup>3</sup>
EI	Índice de emisiones	g/Kg
FAA	Administración Federal de Aviación	-
E10	Mezcla de 10 % de aceite de palma y 90 % de Jet A1	
E20	Mezcla de 20 % de aceite de palma y 80 % de Jet A1	
E50	Mezcla de 50 % de aceite de palma y 50 % de Jet A1	
IATA	Asociación Internacional de Transporte Aéreo	-
NO <sub>x</sub>	Óxidos Nitrosos	-
Oh	Número de Ohnesorge	-
P	Presión	Pa
Pc	Poder calorífico	J/g
Re	Numero de Reynolds	

Nomenclatura	Significado	Unidad
SFC	Consumo específico de combustible	Kg/h
T	Empuje	N
T flama	Temperatura de la flama	K
$W_e$	Índice de Wobbe	J
$\mu_m$	Viscosidad dinámica mezcla	Cst
$\dot{m}_e$	Flujo másico de aire de salida	Kg/s
$V_e$	Velocidad de salida	m/s
$\dot{m}_o$	Flujo másico de aire de entrada	Kg/s
$V_o$	Velocidad de entrada	m/s
$\dot{m}_c$	Flujo másico combinado	Kg/s
L	Longitud de la cámara de combustión	m
$\rho$	Densidad	kg/m <sup>3</sup>
$\eta_{ec}$	Eficiencia de la combustión	
$\dot{G}_{hora}$	Gasto hora	Kg/h
$V_{cc}$	Velocidad cámara de combustión	m/s



## Instrucciones para los Autores

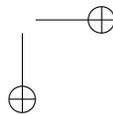
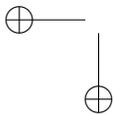
### Introducción

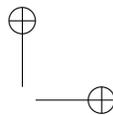
La revista INGENIERÍA es una publicación periódica-científica de la Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas. Tiene como propósito principal divulgar los resultados de investigación de la comunidad académica de la Facultad del país y del exterior. Por tal motivo, los artículos que se publican en la revista deben tener relación con el trabajo en ingeniería, en general, y con las derivaciones de la ingeniería que se desarrollan en la Universidad, en particular. Su frecuencia es cuatrimestral (tres ediciones por año). Todo artículo es evaluado de manera doble-ciega por pares revisores antes de ser calificado como candidato para publicar. El Comité Editorial, tomando en cuenta los conceptos emitidos por los evaluadores y el cumplimiento de las políticas editoriales de la Revista, decide la publicación del artículo en una edición futura. Todos los artículos presentados a la revista INGENIERÍA deben ser inéditos. No se publicarán artículos que hayan sido anteriormente publicados o que hubieren sido sometidos por el autor o los autores a otros medios de difusión nacional o internacional para su publicación.

### Clasificación de los artículos

La clasificación de los artículos dentro de la revista INGENIERÍA se hace de acuerdo con la categorización sugerida por el Departamento Administrativo de Ciencia, Tecnología y e Innovación COLCIENCIAS, para revistas científicas y tecnológicas indexadas. Los autores deben definir y comunicar en la carta de presentación y en la plantilla del artículo la categoría del mismo. Las categorías de los artículos están identificadas por el número de orden de la siguiente lista y sus características son:

1. Artículo de investigación científica y tecnológica: Documento que presenta, de manera detallada, los resultados originales de proyectos de investigación. La estructura generalmente utilizada contiene cuatro apartes importantes: introducción, metodología, resultados y conclusiones.
2. Artículo de reflexión: Documento que presenta resultados de investigación desde una perspectiva analítica, interpretativa o crítica del autor, sobre un tema específico, recurriendo a fuentes originales.
3. Artículo de revisión: Documento resultado de una investigación donde se analizan, sistematizan e integran resultados de investigaciones publicadas o no publicadas, sobre un campo en ciencia o tecnología, con el fin de dar cuenta de los avances y las tendencias de desarrollo. Se caracteriza por presentar una cuidadosa revisión bibliográfica de por lo menos 50 referencias.
4. Artículo cortó: Documento breve que presenta resultados originales preliminares o parciales de una investigación científica o tecnológica, que por lo general requieren de una pronta difusión.
5. Reporte de caso: Documento que presenta los resultados de un estudio sobre una situación particular con el fin de dar a conocer las experiencias técnicas y metodológicas consideradas en un caso específico. Incluye una revisión sistemática comentada de la literatura sobre casos análogos.





6. Revisión de tema: Documento resultado de la revisión crítica de la literatura sobre un tema en particular.

La Revista INGENIERÍA dará prelación a aquellos documentos que se clasificarían en los artículos tipo 1, 2 ó 3.

## **Procedimiento para el envío de artículos para publicación**

Los autores deberán enviar al Editor los siguientes documentos mediante la plataforma Open Journal System (OJS) de la Revista:

- Carta de presentación y originalidad del artículo: El formato de la presente carta puede descargarse en el siguiente enlace:  
<http://revistas.udistrital.edu.co/ojs/index.php/reving/pages/view/Formatos>.  
La finalidad de la carta es certificar que el artículo es inédito, que no está sometido a otra publicación y que se aceptan las normas de publicación de la Revista.
- El artículo en formato digital utilizando la plantilla y las normas de presentación que se pueden consultar en el siguiente enlace (instrucciones a los autores):  
<http://revistas.udistrital.edu.co/ojs/index.php/reving/pages/view/Instructivos>.  
Los autores relacionarán en el correo de postulación del artículo dos posibles evaluadores que cumplan las siguientes condiciones:
  - Debe tener al menos título de maestría
  - Pertenecer a una institución diferente a la de los autores
  - Haber realizado publicaciones en los últimos tres años

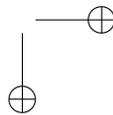
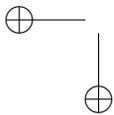
Nota: Enviar los datos de los contactos de los posibles evaluadores (nombre completo, institución y correo electrónico). Esta información también puede ser enviada a la siguiente dirección electrónica [revista\\_ing@udistrital.edu.co](mailto:revista_ing@udistrital.edu.co)

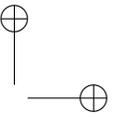
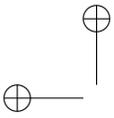
- En caso de ser aceptado, los autores del artículo se comprometen a diligenciar el formato de cesión de derechos para publicación. Se puede descargar en el siguiente enlace:  
<http://revistas.udistrital.edu.co/ojs/index.php/reving/pages/view/Formatos>

Se puede consultar la guía rápida para autores en la plataforma OJS en el enlace:  
<http://revistas.udistrital.edu.co/ojs/index.php/reving/article/view/5540/7070>

## **Aspectos a tener en cuenta**

- Se considera un artículo largo cuando tiene más de diez (10) páginas con este formato. Por consiguiente se recomienda tener como límite diez (10) páginas para la elaboración de un artículo. El Comité Editorial podría solicitar a los autores reducir el tamaño del artículo cuando lo considere demasiado largo.
- Los resultados de las evaluaciones de los pares árbitros y las modificaciones que requiera el Comité Editorial serán notificados a los autores a través de la plataforma OJS y al correo electrónico del autor de correspondencia. Los autores deberán enviar nuevamente los artículos con las modificaciones sugeridas sin nombres y sin biografías. En general, este intercambio de sugerencias y modificaciones puede llegar a presentarse una o dos veces después de presentar el artículo.





- Cuando un artículo no cumple con las condiciones definidas para la revista, el Comité Editorial podría decidir no publicarlo. En este caso, el autor o los autores serán notificados de la decisión.

## **DERECHOS DE AUTOR**

El contenido completo de la licencia Creative Commons, bajo la cual se resguardan los derechos de autor de aquellos que publican en la Revista Ingeniería, puede consultarse en: Creative Commons Attribution-NonCommercial 3.0 Unported License.

